

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS
Departamento de Zoología y Antropología Física



TESIS DOCTORAL

**Estudio de asociación entre marcadores genéticos y alteraciones del
metabolismo del hierro en población española**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Carlos Ignacio Baeza Richer

Directores

Eduardo Arroyo Pardo

Ana María López Parra

Madrid, 2016

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS
Departamento de Zoología y Antropología Física



**Estudio de asociación entre marcadores
genéticos y alteraciones del metabolismo
del hierro en población española**

CARLOS IGNACIO BAEZA RICHER

Tesis Doctoral, 2015

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS
Departamento de Zoología y Antropología Física



**Estudio de asociación entre marcadores
genéticos y alteraciones del metabolismo
del hierro en población española**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA POR

CARLOS IGNACIO BAEZA RICHER

BAJO LA DIRECCIÓN DE LOS DOCTORES

EDUARDO ARROYO PARDO
Dpto. Toxicología y Legislación Sanitaria
Facultad de Medicina
Universidad Complutense de Madrid

ANA MARÍA LÓPEZ PARRA
Dpto. Toxicología y Legislación Sanitaria
Facultad de Medicina
Universidad Complutense de Madrid

Tesis Doctoral
Madrid, 2015

EDUARDO ARROYO PARDO Y ANA MARÍA LÓPEZ PARRA,
DOCTORES Y PROFESORES DEL DEPARTAMENTO DE TOXICOLOGÍA Y
LEGISLACIÓN SANITARIA DE LA FACULTAD DE MEDICINA DE LA
UNIVERSIDAD COMPLUTENSE DE MADRID,

CERTIFICAN QUE:

La Memoria titulada **“Estudio de asociación entre marcadores genéticos y alteraciones del metabolismo del hierro en población española”** que presenta D. **Carlos Ignacio Baeza Richer** para optar al Título de Doctor en Biología, ha sido realizada bajo nuestra dirección. Autorizamos su presentación a la Junta de Facultad de Biología, Departamento de Zoología y Antropología Física de la Universidad Complutense de Madrid al considerar que constituye un trabajo de Tesis.

Para que así conste, firmamos este certificado en Madrid a 29 de septiembre de 2015.

Fdo. Eduardo Arroyo Pardo

Fdo. Ana María López Parra

FINANCIACIÓN

El trabajo de investigación presentado en esta tesis ha sido financiado gracias los proyectos:

- Variabilidad genética en la población española de factores que intervienen en el metabolismo del hierro, FMM-05. Fundación Mutua Madrileña.
- Análisis de polimorfismos genéticos autosómicos, de cromosoma Y y de ADN mitocondrial en muestras biológicas, N8/2006-4130063. Contrato de muestras según el artículo 83 de la LOU. Fundación General de la Universidad Complutense de Madrid.
- Red Temática de Investigación Cardiovascular RECAVA-RD06/0014, PI-08/0420 y PI-08/0756. Fondo de Investigación Sanitaria del Instituto de Salud Carlos III.
- Interacción entre los factores nutricionales y genéticos en la etiología de la ferropenia. Respuesta a un alimento funcional, AGL2009-11437. Ministerio de Ciencia e Innovación.
- Identificación y tipaje de SNPs relacionados con parámetros bioquímicos y hematológicos involucrados en el metabolismo del hierro. Programa de Financiación de Grupos de Investigación Validados-Emergentes, GR35/10. Universidad Complutense de Madrid.

AGRADECIMIENTOS

Cerca ya del final me gustaría agradecer a todos aquellos que han hecho posible este trabajo.

A mis directores, al Prof. Dr. Eduardo Arroyo Pardo y a la Prof. Dra. Ana María López Parra. A Eduardo por haberme tratado siempre como a un igual, por su cercanía, su curiosidad científica y por habernos sentado a todos en la Mesa Redonda. A Ana por su capacidad de trabajo, su instinto científico, paciencia y su inmensa generosidad. A los dos, por toda la ayuda, las correcciones y los consejos, por haber creído en mí aún en los momentos de flaqueza, y por haber arriesgado mezclar trabajo, compañerismo y amistad. Gracias de corazón.

A la Prof. Dra. María Soledad Mesa, mi directora del Diploma de Estudios Avanzados (DEA). A Marisol, por haber guiado mis primeros pasos y abrirme las puertas a la investigación. Por descubrirme la genética de poblaciones. Por su confianza, dedicación y amistad.

Al departamento de Toxicología y Legislación Sanitaria de la facultad de Medicina. Por hacer posible mi trabajo de investigación. A todo el personal docente y al PAS. En especial, al Profesor José María Ruiz de la Cuesta por habernos llenado de libros y sabiduría, y por habernos amenizado con sus “casos bonitos” y con sus conversaciones filosóficas.

A todos mis compañeros de laboratorio, a los que lo fueron y a los presentes. A Mirian, Pedro, Rodrigo, María, Cristina, Eva, Jorge. A Sara, Cláudia y César. Por todas las horas compartidas en nuestra “segunda casa”, el trabajo en equipo, los controles de calidad interminables... Y cómo no, por las risas, las canciones míticas y un largo etcétera de buenos momentos. Habéis sido clave en esta tesis.

Al departamento de Zoología y Antropología Física de la facultad de Ciencias Biológicas, en especial a la unidad de Antropología, por haberme hecho sentir siempre en casa. A los profesores y al personal no docente.

A mis compañeros y amigos de Antropología. A Teresa, Susana, Verónica, Pedro, Labib, Rocío, Jorge, Bea, Gloria, Candela. Por ser compañeros de fatigas, por el conocimiento y la diversión compartida en los congresos. A los doctores y a los doctores en ciernes, estamos cerca.

A la Dra. Pilar Vaquero directora del grupo de Minerales en Metabolismo y Nutrición Humana, y a los colaboradores del ICTAN, Ana, Santiago, Ruth y Laura. A Pilar por haber hecho posible GENUTREN, por su colaboración entusiasta y sus consejos. A todos por el trabajo riguroso y por su labor esencial en el reclutamiento de voluntarios para este estudio y por la recogida de muestras biológicas.

A la Dra. Ana Villegas, anterior Jefa del servicio de Hematología y Hemoterapia del Hospital Clínico de San Carlos, por haber hecho posible la investigación en hemocromatosis. A los colaboradores del Hospital Clínico, a la Dra. Paloma Roperio y al Dr. Fernando Ataúlfo González por su trabajo y buena disposición.

Al Dr. José Manuel Soria, Director de la Unidad de Genómica de Enfermedades Complejas del Instituto de Investigaciones Biomédicas del Hospital Sant Pau, por su colaboración y por haberme acogido en mi estancia en Barcelona. A los doctores Alfonso Buil, Juan Carlos Souto y Anna Brichs.

Al Dr. Ángel Remacha facultativo del Servicio de Laboratorio de Hematología del Hospital Sant Pau de Barcelona. Por su colaboración, por la cesión de muestras y por confiar en mi criterio científico.

A la Dra. María del Carmen Bravo del Servicio de Apoyo al Usuario-Investigación en la UCM, por su inestimable ayuda en los aspectos estadísticos de esta tesis.

A todos mis amigos. A los de la infancia en Rota, ese dulce y nostálgico tercio de mi vida. A mis profesores, desde el colegio a la facultad, a los que me enseñaron el valor de conocer y dar a conocer, y a los que despertaron mi pasión por la biología, espero saber transmitírsela a mis ahijados y sobrinos. A mis compañeros y amigos de la facultad. A Esther y Rubén por animarme continuamente. Al “Biogrupo” por los buenos momentos y los viajes en “comuna”. A mis amigos de Aranjuez, a Carlos y a Javi por su amistad de tantos años. A mis amigos los teatreros de “El Submarino”, es una suerte tenerlos. En especial a Jo y a Diego por su ayuda gráfica.

A mi prima Rocío, por sus revisiones de textos en inglés, por su apoyo incondicional y por ser uno de los pilares en mi vida. A mi familia extensa, la Baeza y la Richer, a mi cuñada Ana, tíos, primos, sobrinos, sois una fuente de cariño y apoyo.

A mis padres y a mi hermano. Las palabras no hacen justicia a todo lo que tengo que agradecerlos. A mis padres por su amor incondicional, sus ánimos y por su confianza ciega. A mi padre por inculcarme la curiosidad, a mi madre por infundirme alegría. A mi hermano por ver en mí más de lo que yo veía y por saber que siempre puedo contar contigo pese a la distancia.

A Vero, a mi niña, mi vida. No tengo espacio ni palabras para darte las gracias como te mereces. Gracias por tu ayuda desinteresada, por darme el empujón adecuado en el momento adecuado. Por prohibirme decir no puedo. Por creer. Por encontrar siempre el lado bueno. Por compartir lo amable y sobre todo lo duro. Gracias por estar, por ser.

A mis Padres e Irmão

A Verónica

*“El problema del físico es el problema de los orígenes y leyes naturales últimas.
El problema del biólogo es el problema de la complejidad”*

Richard Dawkins

*“Antes pensábamos que nuestro futuro estaba en las estrellas.
Ahora sabemos que está en nuestros genes”*

James Watson

“No está en los genes”

Richard C. Lewontin

“Soy de los que piensan que la ciencia tiene una gran belleza. Un sabio en su laboratorio no es solamente un teórico. Es también un niño colocado ante los fenómenos naturales que le impresionan como un cuento de hadas”

Marie Curie

Esta tesis doctoral corresponde a la modalidad de compendio de publicaciones, incluyendo los cinco trabajos siguientes*:

- Bertoncini S, Blanco-Rojo R, **Baeza C**, Arroyo-Pardo E, Vaquero MP, López-Parra AM. A novel SNaPshot assay to detect genetic mutations related to iron metabolism. *Genet Test Mol Biomarkers* 2011;15(3):173-9.
- **Baeza Richer C**, López-Parra AM, Blanco-Rojo R, Bertoncini S., Arroyo-Pardo E, Vaquero MP. Variabilidad genética en estudios de asociación. Estructura poblacional de 10 SNPs autosómicos relacionados con el metabolismo del hierro en mujeres españolas. En: *Biodiversidad Humana y Evolución* 2012. Estudi editorial. Diversidad Genética de las Poblaciones Humanas (Cap. 5): 413-6.
- **Baeza-Richer C**, Roperio P, González FA, Mesa MS, López-Parra AM, Villegas A, M.P. Vaquero MP, Arroyo-Pardo E. HFE, Hemojuvelin and Hpcidine genotyping confirms high prevalence of non-HFE iron overload in Southern European population. (Bajo consideración editorial en *Hum Biol*).
- **Baeza-Richer C**, Blanco-Rojo R, López-Parra AM, Brichs A, Bertoncini S, Pérez-Granados AM, Buil A, Soria JM, Arroyo-Pardo E, Vaquero MP. Identification of a novel Quantitative Trait Nucleotype related to iron status in a calcium channel gene. *Dis Markers* 2013;34(2):121-9.
- **Baeza-Richer C**, López-Parra AM, Blanco-Rojo R, Toxqui L, Remacha A, Vaquero MP, Arroyo-Pardo E. Genetic contribution to iron status: SNPs related to iron deficiency anaemia and fine mapping of CACNA2D3 calcium channel subunit. *Blood Cells Mols and Dis* 2015;55(4):273-80.

*Además, se incluye un sexto trabajo, *no publicado*, que fue presentado en el congreso internacional *Portugalix Genetica, Instituto de Patología Molecular e Inmunología de la Universidad de Oporto (IPATIMUP)*, Portugal en 2006. Resultados Complementarios: *Frequency distribution and detection of mutations in genes implicated in iron deficiency in women*.

LISTADO DE ABREVIATURAS

ADN	Ácido DesoxirriboNucleico
ANOVA	Análisis de la varianza (<i>ANalysis Of VAriance</i>)
BMA	Promediado bayesiano de modelos (<i>Bayesian Model Averaging</i>)
BMP	Proteína morfogenética (<i>Bone Morphogenetic Protein</i>)
BMPR	Receptor de proteína morfogenética (<i>Bone Morphogenetic Protein Receptor</i>)
CDC	<i>Centers for Disease Control and Prevention, USA</i>
CEGEN	Centro Nacional de Genotipado, España
CEPH	Muestra del proyecto <i>HapMap</i> comprendida por individuos residentes en Utah descendientes del norte y oeste de Europa
CHAID	Detección automática de interacciones mediante chi-cuadrado (<i>Chi-square Automatic Interactions Detection</i>)
DMT1	Proteína transportadora de metales divalentes 1 (<i>Divalent Metal Transporter 1</i>)
Dyctb	Citocromo duodenal B (<i>Duodenal cytochrome B</i>)
FPN	Proteína y <i>gen</i> (en cursiva) de la ferroportina
Ft	Ferritina
GWAS	Estudio de asociación mediante mediante genotipado masivo (<i>Genome Wide Association Study</i>)
HAMP	Proteína y <i>gen</i> (en cursiva) de la Hepcidina
HapMap	Mapa de haplotipos (<i>Haplotype Map</i>)
Hb	Hemoglobina
HCP1	Proteína transportadora de hemo 1 (<i>Heme Carrier Protein 1</i>)
Hct/Ht/PCV	Hematocrito (<i>Pack Cell Volume</i>)
HFE	Proteína y <i>gen</i> (en cursiva) de la hemocromatosis humana (<i>High Fe</i>)
HH	Hemocromatosis Hereditaria

HIF	Factores de transducción inducidos por hipoxia (<i>Hypoxia Inducible Factor</i>)
HJV	Proteína y <i>gen</i> (en cursiva) de la Hemojuvelina
HWE	Equilibrio Hardy-Weinberg (<i>Hardy-Weinberg Equilibrium</i>)
IDA	Anemia por deficiencia de hierro (<i>Iron Deficiency Anaemia</i>)
IL-6	Interleucina-6
IRE	Elementos sensibles al hierro (<i>Iron Resposive Elements</i>)
LD	Desequilibrio de ligamiento (<i>Linkage Disequilibrium</i>)
MAF	Frecuencia de alelo derivado (<i>Minimum Allele Frequency</i>)
NHGRI	<i>National Human Genome Research Institute, USA</i>
OMIM	Catálogo de enfermedades humanas de herencia mendeliana (<i>Online Mendelian Inheritance in Man</i>)
OMS/WHO	Organización Mundial de la Salud (<i>World Health Organization</i>)
PCR	Reacción en cadena de la polimerasa (<i>Polymerase Chain Reaction</i>)
RFLP	Polimorfismo de la longitud de los fragmentos de restricción (<i>Restriction Fragment Lenght Polymorphism</i>)
RFU	Unidades relativas de fluorescencia (<i>Relative Fluorescence Units</i>)
rs#####	Número de referencia de SNP (<i>Reference SNP number</i>)
SAT	Saturación de la transferrina
SMAD	Proteína mediadora de señales intracelulares
SNP	Polimorfismo de un solo nucleótido (<i>Single Nucleotide Polymorphism</i>)
SSCP	Polimorfismo de conformación de la cadena sencilla (<i>Single Strand Conformation Polymorphism</i>)
STAT	Proteína transductora de señal y activadora de transcripción
STR	secuencias cortas repetidas en tándem (<i>Short Tandem Repeat</i>)
TF	Gen de la transferrina

Tf	Transferrina
TfR	Receptor de transferrina
TMPRSS6	Proteína y <i>gen</i> (en cursiva) de la matriptasa 2
VCM	Volumen Corpuscular Medio

ÍNDICE

RESUMEN.....	21
Resumen	23
Summary	27
1. INTRODUCCIÓN.....	31
1.1. Metabolismo del hierro	33
1.1.1. El hierro y la importancia de su regulación	33
1.1.2. Absorción, transporte y reciclaje del hierro	34
1.1.3. Detección y regulación del hierro sistémico	36
1.2. Evaluación del estado del hierro	39
1.3. Trastornos del metabolismo del hierro	41
1.3.1. Anemia ferropénica	41
1.3.2. Otros tipos de anemia	42
1.3.3. Hemocromatosis	43
1.4. Marcadores genéticos relacionados con el metabolismo del hierro.....	47
1.4.1. Heredabilidad y desequilibrio de ligamiento	47
1.4.2. Variabilidad poblacional y proyecto <i>HapMap</i>	49
1.4.3. Estudios de asociación y base genética de las alteraciones del metabolismo del hierro	51
1.5. Objetivos	55
2. DESARROLLO DE UNA HERRAMIENTA DE GENOTIPADO PARA DETECTAR MARCADORES GENÉTICOS RELACIONADOS CON EL METABOLISMO DEL HIERRO.....	57
2.1. “A novel SNaPshot assay to detect genetic mutations related to iron metabolism”	59
3. VARIABILIDAD DE LAS MUTACIONES RELACIONADAS CON EL METABOLISMO DEL HIERRO EN POBLACIÓN ESPAÑOLA.....	77
3.1. “Variabilidad genética en estudios de asociación. Estructura poblacional de 10 SNPs autosómicos relacionados con el metabolismo del hierro en mujeres españolas”	79
3.2. “HFE, Hemojuvelin and Hepcidine genotyping confirms high prevalence of non-HFE iron overload in Southern European population”	89

4. ESTUDIOS DE ASOCIACIÓN DE MARCADORES GENÉTICOS CON PARÁMETROS BIOQUÍMICOS Y FENOTIPOS CLÍNICOS RELACIONADOS CON EL METABOLISMO DEL HIERRO	105
4.1. "Identification of a novel Quantitative Trait Nucleotype related to iron status in a calcium channel gene"	107
4.2. "Genetic contribution to iron status: SNPs related to iron deficiency anaemia and fine mapping of CACNA2D3 calcium channel subunit"	131
4.3. "Frequency distribution and detection of mutations in genes implicated in iron deficiency in women"	161
5. DISCUSIÓN	171
5.1. Discusión por objetivos	174
5.2. Temas transversales y consideraciones generales	193
6. CONCLUSIONES.....	201
7. BIBLIOGRAFÍA.....	209

RESUMEN

Resumen

El hierro es un elemento esencial para numerosos procesos fisiológicos. Forma parte de la hemoglobina, proteína encargada de administrar oxígeno a los tejidos. También interviene en otros procesos como la respiración celular a nivel de la mitocondria y se integra en otras proteínas como la mioglobina. Sin embargo, un exceso de hierro puede causar daños en los tejidos al generarse radicales libres altamente tóxicos, mientras que una carencia puede derivar en anemia. Los trabajos incluidos en esta tesis doctoral se centran, preferentemente, en la anemia ferropénica. Este tipo de anemia es la más frecuente, siendo un trastorno de prevalencia mundial. Su herencia es compleja, ya que múltiples factores ambientales y genéticos están involucrados en su etiología. Por otro lado, se estudia la hemocromatosis, una patología causada por los daños tisulares debidos al exceso de hierro. La base genética de su variante más frecuente está bien descrita y consiste en la mutación C282Y del gen *HFE*. Sin embargo, su penetrancia incompleta y la relativa baja frecuencia de esta mutación en poblaciones mediterráneas hacen necesario la búsqueda de otras mutaciones, que expliquen mejor la etiología y los diferentes grados de gravedad de esta enfermedad. Además los valores relativamente altos de heredabilidad de los principales parámetros bioquímicos y hematológicos empleados para valorar el estado del hierro (hemoglobina, ferritina, hematocrito, hierro sérico, volumen corpuscular medio, transferrina y saturación de la transferrina) justifican la búsqueda de su base genética, y la de las alteraciones del metabolismo del hierro.

La presente tesis doctoral tiene como objetivo general encontrar marcadores genéticos asociados a las alteraciones del metabolismo del hierro, concretamente a la anemia ferropénica y a la hemocromatosis. Los objetivos específicos son:

1. Diseñar y valorar una herramienta de genotipado (reacción de mini-secuenciación *SNaPshot*), para replicar la asociación entre marcadores genéticos y parámetros bioquímicos relacionados con el metabolismo del hierro.

2. Comparar la variabilidad de los marcadores incluidos en el *SNaPshot* entre población española y otras caucásicas. Estudiar su posible estratificación poblacional, con el fin de valorar la utilidad clínica de estas variantes.
3. Buscar variantes en los genes candidato del exceso de hierro y/o hemocromatosis (hepcidina y hemojuvelina) en individuos con diferentes genotipos para el gen *HFE*. Comparar su distribución alélica con la de otras poblaciones.
4. Establecer asociaciones entre marcadores genéticos y parámetros bioquímicos y hematológicos relacionados con el metabolismo del hierro.
5. Estudiar las asociaciones obtenidas más relevantes, con el fin de proponer nuevas hipótesis sobre su implicación en la etiología de la anemia ferropénica.
6. Establecer un modelo para identificar factores genéticos de riesgo o protección que influyen en el desarrollo de la anemia ferropénica.

Para ello, en primer lugar se desarrolló una herramienta de genotipado por *SNaPshot* para analizar, en una sola reacción, 10 marcadores o polimorfismos de un único nucleótido (SNPs) seleccionados por su posible asociación con el estado del hierro. En segundo lugar, se realizaron dos estudios genético-poblacionales en los que se analizaron marcadores o SNPs relacionados con anemia ferropénica y hemocromatosis en población española y se establecieron comparaciones con las frecuencias encontradas en otras poblaciones europeas. En tercer y último lugar, se realizaron tres estudios de genética clínica sobre la anemia ferropénica, utilizando el *SNaPshot* desarrollado y añadiendo nuevos SNPs seleccionados de la literatura. En total, se estudiaron hasta 45 SNPs en 358 mujeres españolas autóctonas en edad fértil.

Los resultados obtenidos en esta tesis doctoral confirman que el sistema de genotipado diseñado es una técnica sencilla, consistente, sensible y con una baja

tasa de error. Además, la selección de los 10 marcadores fue correcta ya que la mayoría se asoció a algún parámetro del estado del hierro. A nivel poblacional, las frecuencias encontradas en población española para los 10 SNPs relacionados con la anemia ferropénica no difieren de las publicadas en otras poblaciones europeas.

En cuanto a la hemocromatosis, se demuestra que los genes hemojuvelina y hepcidina están muy conservados en individuos de población española con diferentes genotipos para el gen *HFE*. Además, se confirma la mayor prevalencia de hemocromatosis no debida a *HFE* en poblaciones del sur de Europa. Los estudios de asociación sobre anemia ferropénica muestran en esta tesis, por primera vez, que el SNP rs1375515 está asociado a los niveles de hemoglobina y ferritina sérica, y es también un factor de riesgo para este tipo de anemia. Este SNP y el rs1800562 explican diferencias significativas en la distribución de los individuos en los tres fenotipos clínicos (normal, deficiente y anémico).

Por otro lado, se han replicado la asociación de varios SNPs como por ejemplo el rs3811647 y el rs1799852 con los niveles de transferrina, o la mutación C282Y (rs1800562) con el estado del hierro, entre otros. Además, se han identificado dos factores de protección de la anemia ferropénica, el rs1800562 (localizado en el gen *HFE*) y el rs4895441 (región intergénica *HB1S-MYB*, relacionado con altos niveles de hemoglobina fetal). Por el contrario, el rs1314491 (gen *HIST1H2BJ*, que codifica una histona tipo 1, y el rs1868505 (gen *CACNA2D3*) son factores de riesgo. La localización de rs1375515 en la subunidad CACNA2D3 de un canal del calcio motivó su análisis mediante *Fine Mapping*. La asociación de los SNPs de *CACNA2D3* con ferritina y hierro sérico sugiere que estos canales de calcio pueden ser una vía de entrada del hierro al interior celular, y que, por tanto, estos marcadores modificarían de alguna forma su funcionalidad.

En conclusión, los hallazgos de la presente tesis doctoral han permitido profundizar en el conocimiento de la asociación entre marcadores genéticos y los trastornos del metabolismo del hierro al haber: a) mostrado como la hepcidina y la hemojuvelina son genes altamente conservados en población española con sobrecarga de hierro, b) replicado, por primera vez, la asociación de determinados marcadores con parámetros relacionados con el estado del hierro y con la anemia

ferropénica en población española, c) identificado un nuevo SNP (rs1375515), asociado con el estado del hierro, d) proporcionado evidencia genética de que los canales de calcio podrían ser una vía de entrada de hierro en condiciones fisiológicas.

Summary

Iron is an essential element for normal organism function, thus, fine regulation of this element is required because departures from its optimal levels produce severe alterations. For example, iron overload produces hydroxyl radicals which cause severe tissue damages. By contrast, iron deficiency may cause anaemia. As mammals have no active mechanisms to eliminate iron, its absorption and storage regulation are crucial to maintain appropriate levels of this mineral.

Not only are nutritional and host-related factors determinants of iron levels, but also genetic factors play an important role. The heritability of the biochemical parameters used to assess iron status, such as haemoglobin, haematocrit, ferritin, mean cell volume, transferrin and serum iron, supports this hypothesis. Several studies on humans have associated many genetic markers, mainly SNPs, with these parameters and with anaemia itself. However, it is still difficult to replicate or apply the obtained results in practical health contexts, rendering it necessary to broaden methodological strategies in research, using data mining methods as a means to integrate all of these associations.

Among iron related diseases, this thesis mainly focuses on iron deficiency anaemia that, being the most worldwide spread, is considered as a pandemic. Iron overload and hemochromatosis have also been studied. The genetic basis of type 1 of this disease is well established and consists in mutations in the *HFE* gene. However, the incomplete penetrance and the relatively low frequency of these mutations in southern European populations makes it necessary to explore new genes that could explain better the etiology of this pathology.

The general aim of this thesis work is to broaden the knowledge about the association between genetic markers and the iron metabolism disorders. The specific aims are:

1. To design and test a genotyping tool (mini-sequencing reaction SNaPshot) in order to replicate the association between genetic markers and biochemical parameters related to iron metabolism.

2. To compare the variability of the markers included in the SNaPhot, in Spanish and other Caucasian populations. To study its possible structure in Spanish population in order to assess its clinical utility.
3. To search for variants in the candidate genes for iron overload and/or hemochromatosis (hepcidin and hemojuvelin) in individuals with different genotypes of the *HFE* gene. To compare their allelic distribution with those of other populations.
4. To establish associations between genetic markers and iron-related biochemical and hematological parameters.
5. To study the most relevant associations obtained, aiming to propose new hypothesis about their possible involvement in the etiology of iron deficiency anaemia.
6. To establish a model to identify genetic risk or protective factors that play a role in developing iron deficiency anaemia.

This thesis includes six works that address the aforementioned objectives.

The first study is a methodological work in which a genotyping tool that targets 10 SNPs in genetic regions related to iron status was developed. The results obtained confirm that the developed mini-sequencing reaction SNaPshot is a very simple, sensitive, low cost and precise method. Furthermore, the majority of the SNPs were found associated with at least one iron related biochemical parameter, which confirms that the SNP selection was appropriate.

In the second study, individuals classified according to the three iron-clinical phenotypic groups (normal deficient and anaemic) were genotyped in order to evaluate the SNaPshot reaction. Genetic distances were assessed among this Spanish sample, as a whole and divided into the three groups, with Caucasian populations available in the NCBI Reference Assembly. No significant genetic distances were found, confirming that the markers selected do not reveal any population substructure. These findings reinforced the possible clinical utility of the developed SNaPshot reaction.

In regard to the hemochromatosis, the third work explores the presence of mutations in hepcidin and hemojuvelin in patients with different genotypes of *HFE* gene (C282Y and H63D) in Spanish individuals suffering from iron overload. The influence of these genotypes on iron related parameters was also assessed. The results showed that the hemojuvelin and hepcidin genes are highly conserved in Spanish population. Moreover, the higher prevalence of non-*HFE* hemochromatosis in southern European populations was confirmed.

In the fourth research, 284 women at fertile age were genotyped using the SNaPshot reaction. The results showed, in this thesis and for the first time, that the SNP rs1375515 is associated with the levels of haemoglobin and serum ferritin and that this genetic marker is a risk factor for developing this kind of anaemia. This SNP along with the rs1800562 explained significant differences in the distribution of individuals in the three iron clinical phenotypes (normal, deficient and anaemic).

The fifth work analyzes 358 Spanish women at fertile age and tested 45 SNPs. This association study was able to replicate the relation between several SNPs, such as rs3811647 and rs1799852, and the levels of serum transferrin, or the association of C282Y mutation with iron status. Moreover, two protection factors regarding iron deficiency anaemia have been identified: the rs1800562 (located in *HFE* gene), and the rs4895441 (located in *HB1SL-MYB*, related with persistent high levels of fetal hemoglobin). On the opposite, rs1314491 (in *HIST1H2BJ*, which codifies a type 1 histone) and rs1868505 (in *CACNA2D3*) were found to be risk factors for this disorder. Furthermore, the rs1375515 placement in the calcium channel subunit, *CACNA2D3*, motivated a Fine Mapping study of this region. The associations found for the SNPs located in this region suggest that the calcium channels could be a direct way for the iron to enter the cells, and that these SNPs could modify, in some manner, their functionality.

In conclusion, the findings of the present thesis work have broadened the knowledge of the genetic and molecular basis of iron related disorders by means of: a) showing how hepcidin and hemojuvelin genes are highly conserved in Spanish population with iron overload, b) replicating, for the first time, the

association of certain genetic markers with iron-related parameters and iron deficiency anaemia in Spanish population, c) identifying a novel SNP, rs1375515, related to iron status, d) providing genetic evidence that calcium channels could be a direct way of entrance for iron in physiological conditions.

1. INTRODUCCIÓN

1.1. Metabolismo del hierro

1.1.1. El hierro y la importancia de su regulación

El hierro es un nutriente esencial para numerosos procesos fisiológicos del organismo. Además, es un componente de los citocromos y de la molécula hemo, que a su vez forma parte de la hemoglobina y la mioglobina (Zhang y Enns 2009). El hierro es un cofactor en enzimas como la ribonucleotido-reductasa, que controla la síntesis de ADN. También juega un papel fundamental en los procesos de oxidoreducción en las cadenas respiratorias de las mitocondrias, mecanismo clave en la respiración celular (Camaschella 2015). En general, participa de la integridad y mantenimiento celular formando parte de catalasas, peroxidasas y oxigenasas (Anderson et al. 2007).

Al contrario que otros elementos químicos esenciales como el zinc, el hierro es un elemento abundante en la naturaleza (Abbaspour et al. 2014). Sin embargo, a pesar de la abundancia geológica, su biodisponibilidad no es tan alta. El hierro genera con facilidad óxidos altamente insolubles que son difíciles de ingerir por los organismos. En respuesta a este problema, algunos seres vivos han desarrollado diversos mecanismos celulares para capturar el hierro en formas biológicamente útiles. Ejemplo de ello sería el mecanismo por el cual la levadura es capaz de reducir la forma insoluble del hierro férrico (Fe^{3+}) a hierro ferroso (Fe^{2+}), fácilmente soluble en agua. Los organismos superiores como el ser humano poseen mecanismos análogos al comentado. En el organismo humano, el hierro se encuentra formando parte de complejos proteínicos como la hemoglobina y la mioglobina (grupo hemo) o en otros complejos no relacionados con el grupo hemo, como en los enzimas hierro-flávicos, en la ferritina y en la transferrina. Prácticamente dos tercios del hierro de todo el organismo se encuentran unidos a la hemoglobina dentro de los glóbulos rojos circulantes, 25% almacenado y 15% unido a la mioglobina en el tejido muscular y a los enzimas ya comentados. El conjunto de la ferritina y la hemosiderina se considera la reserva corporal de este metal (Hentze et al. 2010).

Las particularidades en el almacenado del hierro hacen que éste sea muy tóxico cuando se acumula de forma no controlada. El exceso de hierro genera radicales hidroxílicos y aniones oxidantes altamente reactivos con moléculas biológicas, como proteínas, lípidos y/o ADN. Como se describirá en el apartado 1.4 el hierro desencadena la reacción Fenton, por la cual se generan radicales libres que son altamente tóxicos para los tejidos (Hentze et al. 2004). Además, el organismo no cuenta con ningún mecanismo activo de secreción y las pérdidas de hierro se dan pasivamente por sudoración, descamación del epitelio digestivo y en la menstruación. Esto hace que el único punto de control se dé en la absorción a nivel de los enterocitos (Abbaspour et al. 2014).

1.1.2. Absorción, transporte y reciclaje de hierro

El hierro se absorbe a partir de la dieta de dos formas: en forma hemo y en forma no hemo. Una de ellas, la forma no hemo, es más biodisponible que la otra. Curiosamente, el organismo absorbe más fácilmente hierro en la forma hemo que es más escasa en la dieta (Hurrell y Egli 2010).

A nivel del duodeno, el hierro se introduce en los enterocitos ya sea por medio del *transportador de metales divalentes 1* (DMT1) en la forma no hemo, o mediante la *proteína transportadora de hemo 1* (HCP1), aunque existe cierta controversia sobre este punto (Zhang y Enns 2009). Para que el DMT1 pueda actuar, el hierro procedente de la dieta debe ser transformado en Fe^{2+} , función de la que se ocupa principalmente la reductasa DcytB. En el citosol de los enterocitos, el hierro se puede almacenar uniéndose a la *ferritina* (Ft) o atravesar la membrana basolateral para ser transportado al resto del organismo. En esta membrana se encuentra la *ferroportina* (FPN), proteína implicada en el transporte y reciclaje del hierro y sobre la cual actúa la *hepcidina*, como se comentará posteriormente. La FPN exporta hierro del interior al espacio extracelular en los enterocitos, macrófagos y hepatocitos. Una vez extraído por la FPN, la *hefastina* oxida el Fe^{2+} a Fe^{3+} para que la *transferrina* (Tf), la proteína más importante en el transporte del hierro (Hentze et al. 2010), haga circular este mineral por todo el organismo en

función de los requerimientos fisiológicos. El hierro se transporta a la médula para la eritropoyesis, al tejido muscular y al interior celular para funciones respiratorias. La mayor parte del hierro empleado para la eritropoyesis proviene del hierro reciclado de los eritrocitos senescentes al ser fagocitados por los macrófagos. Este proceso se da mayoritariamente en el hígado, que es el órgano más importante en la regulación de la homeostasis del hierro. Se calcula que de 1 a 2 mg de hierro se introducen por la dieta, mientras que se reciclan de 20 a 25 mg (Figura 1). El receptor de transferrina 1 (TfR1), localizado en la membrana de la mayoría de los tipos celulares, es el encargado de captar la transferrina e introducir el hierro al interior citoplasmático.

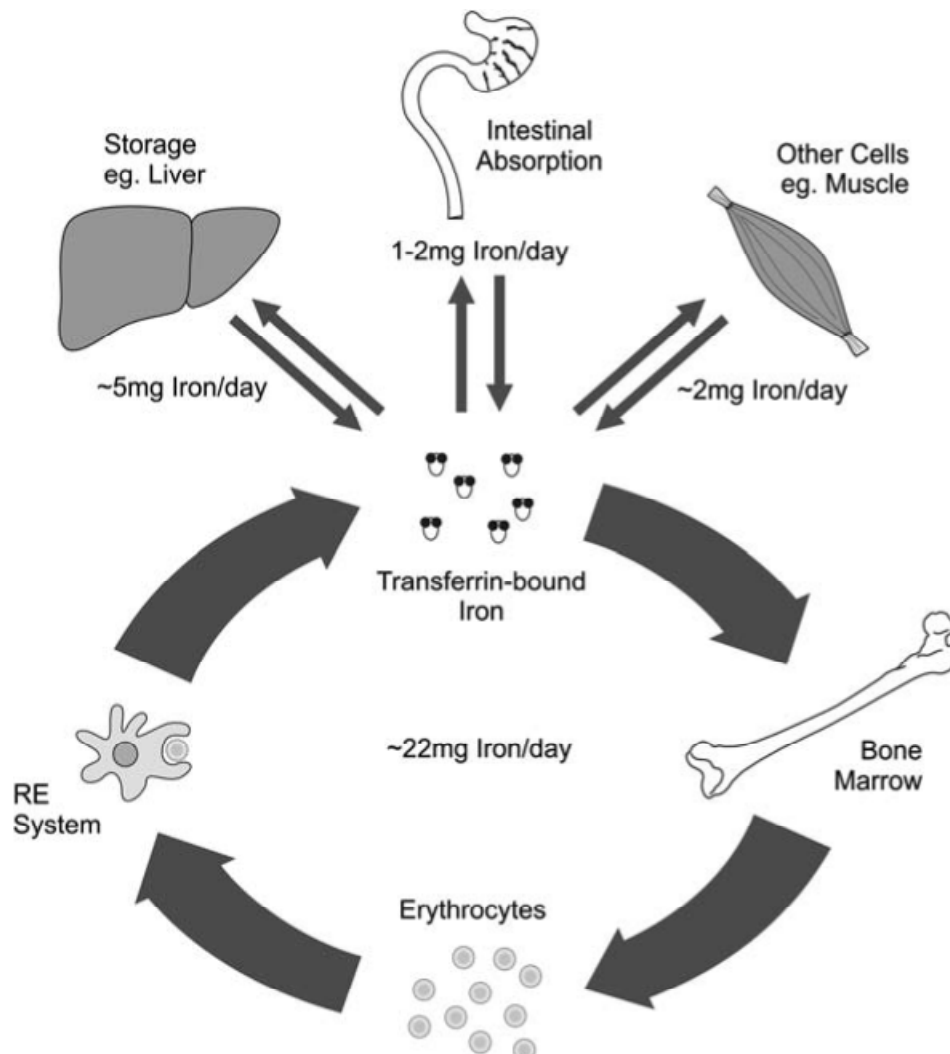


Figura 1. Absorción, transporte y reciclaje de hierro sistémico. Se indica el flujo de hierro (mg/día) en cada uno de estos procesos. Tomado de Anderson et al. 2007

1.1.3. Detección y regulación del hierro sistémico

Como se comentó anteriormente, el hígado es el órgano encargado de la regulación de los niveles sistémicos de hierro. En él se secreta la *hepcidina*, considerada la proteína clave en la homeostasia del hierro (Abbaspour et al. 2014). Todos los procesos implicados en la regulación de este mineral, ya sean las variaciones en los niveles de hierro, la hipoxia, la inflamación y la eritropoyesis influyen en la expresión de la hepcidina inhibiéndola o estimulándola (Figura 2). La hepcidina es una proteína anti-microbiana mediadora de los procesos inmunitarios (Park et al. 2001). En condiciones de hierro elevado, esta proteína se une a la FPN y produce su endocitosis y posterior degradación. Esto provoca que el hierro no pueda exportarse del interior celular, impidiendo su absorción a nivel de los enterocitos, el reciclaje en los macrófagos y la liberación del hierro almacenado

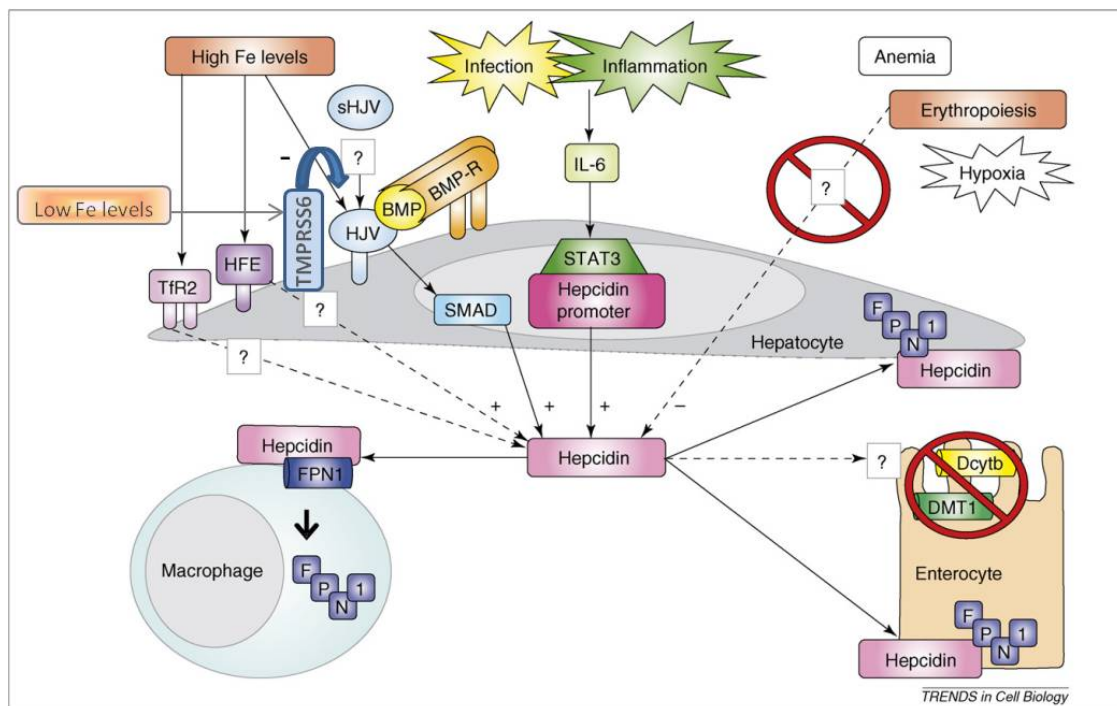


Figura 2. Regulación de los niveles de hierro sistémico a nivel del hepatocito. La hepcidina se muestra como el eje hacia el que converge toda la red de señalización. La hepcidina provoca la degradación de la ferroportina (FPN) con lo que el hierro deja de absorberse en el duodeno (enterocitos), impidiendo además su recirculación sistémica. La matriptasa 2 (TMPS6) degrada la hemojuvelina (HJV) en condiciones de hierro bajo. IL6: Interleukina-6. STAT3: Transductor de señal y activador de la transcripción 3. DMT1: Transportador de metales divalentes 1. Dcytb: Reductasa. SMAD: Mediador de señal intracelular. Modificado de Dunn et al. 2007

en la ferritina. Cuando hay bajos niveles de hierro se inhibe la expresión de la hepcidina, con lo que se invierten los procesos anteriores al estar activa la FPN. Es decir, se absorbe hierro, los macrófagos liberan el hierro captado de los eritrocitos senescentes, y se libera el almacenado. Esta es la razón por la cual una sobreexpresión de la hepcidina produciría anemia, y su ausencia, sobrecarga de hierro (Hentze et al. 2010). Mediante el control de los niveles de hepcidina se regula el hierro sistémico, hallándose en el hígado un complejo de proteínas que se encargan de controlar su expresión, como son el HFE, las citokinas BMPs (proteínas morfogenéticas óseas), la hemojuvelina (HJV), la matriptasa 2 (TMPRSS6) y el receptor de transferrina 2 (TfR2), entre otras (Figura 2).

La *hemojuvelina* (HJV) se expresa mayoritariamente en músculo esquelético, corazón y, en menor medida, en el hígado. Es una glicoproteína anclada a la membrana plasmática que funciona como receptor de las BMPs, estimulando la expresión de la hepcidina a través de una cascada de señalización, en la que intervienen las proteínas de transducción SMADs y que finaliza en un aumento de la expresión de la hepcidina a nivel nuclear (Figuras 2 y 3). La forma soluble sHJV se une a la anclada en la membrana, interrumpiendo la estimulación de la expresión de hepcidina. Ciertos trabajos señalan que la expresión de HJV soluble es dependiente del estado del hierro y se inhibe como respuesta a los niveles de Tf cargada de hierro (Zhang et al. 2007, Core et al. 2014). Al estar implicada en varios sistemas de regulación, algunos autores han considerado el sistema HJV-hepcidina como el eje regulador del hierro (Zhang 2010).

El *receptor de transferrina 2* (TfR2), tiene una afinidad por la Tf menor que el TfR1 y se ha sugerido que está más implicado en generar una señal de aumento de la expresión de hepcidina que en interiorizar hierro al citoplasma (Silvestri et al. 2014). Por tanto, se le considera uno de los sensores del estado del hierro (Figuras 2 y 3). Es, además, una proteína poco estable, lo que favorecería los cambios rápidos en concentración a nivel de la membrana de los hepatocitos.

La *HFE* es una proteína transmembrana del complejo mayor de histocompatibilidad. Localizada en la superficie de los hepatocitos (Figura 3), funciona también a modo de sensor de los niveles de hierro e interacciona con el

TfR1, el TfR2 y la transferrina (Hentze et al. 2010, Gkouvatsos et al. 2012). En condiciones de hierro alto, el HFE forma un complejo con el TfR2 que emite una señal de transducción para que se exprese la hepcidina. Otros trabajos han demostrado también como el HFE estimula la expresión del mismo TfR2, lo que refuerza su papel regulador.

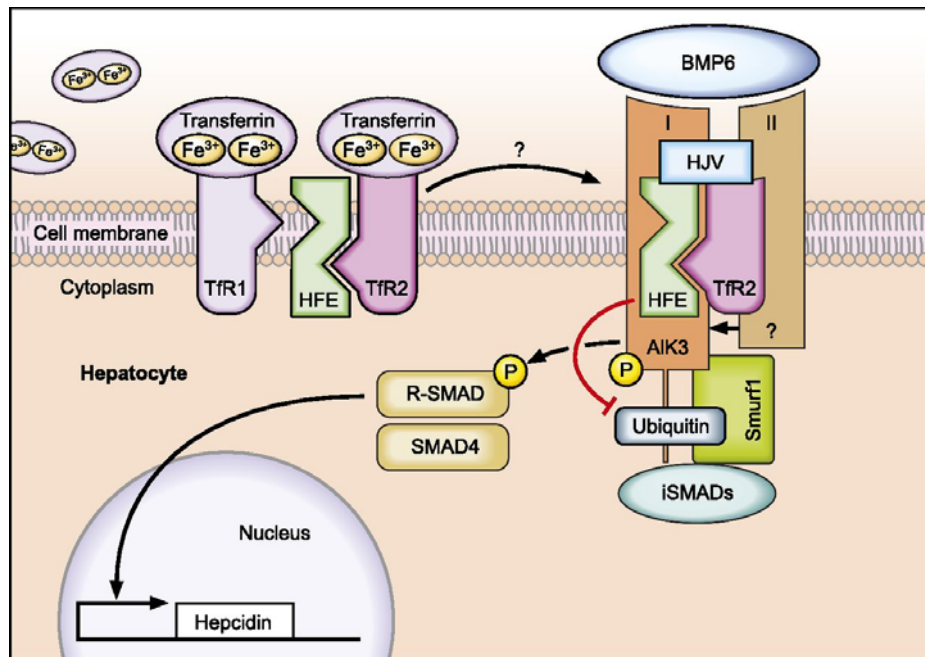


Figura 3. Cascada para estimular la expresión de hepcidina a nivel nuclear. En condiciones de hierro alto, (transferrina con sus dos sitios de unión al hierro ocupados), el HFE se desacopla del receptor de transferrina 1 (TfR1) y forma un complejo con la hemojuvelina (HJV), el receptor de transferrina 2 (TfR2), la proteína morfogenética 6 (BMP6) y sus receptores BMPRI (también llamado Alk3) y II. Este complejo emite una señal de transducción para inducir la síntesis de hepcidina a través de las proteínas mediadoras de señales intracelulares SMAD. Tomado de Muckenthaler 2014

La *matriptasa 2* (TMPRSS6) es una proteína transmembrana con función serin proteasa, implicada en la regulación, pero con efecto opuesto al de HFE. Por medio de mecanismos aún discutidos, parece detectar niveles bajos de hierro mediando una señal de inhibición de la expresión de hepcidina, lo que generaría un aumento de absorción a nivel entérico. Se ha sugerido que la TMPRSS6 produce la

proteólisis de la HJV (Figura 2) que, al inactivarse, dejaría de estimular la secreción de hepcidina (Silvestri et al. 2008, Wang et al. 2014).

Los procesos infecciosos, inflamatorios y la hipoxia, poseen mecanismos, independientes a los ya comentados, que influyen en la expresión de hepcidina (Figura 2). La hipoxia dispara la expresión de eritropoyetina y esto desencadena, por un lado, la eritropoyesis y, por otro, la inhibición de la hepcidina a través de los factores de transducción inducidos por hipoxia o HIFs (Dunn et al. 2007). Los HIFs se translocan al núcleo y se unen al promotor de la hepcidina impidiendo que ésta se exprese (Hentze et al. 2010). En los procesos infecciosos e inflamatorios se potencia la expresión de esta proteína, independientemente de los niveles de hierro, a través de la interleukina 6, que a su vez activa los factores de transcripción STAT3. Este mecanismo es la base de la anemia por enfermedad crónica (Zhang y Enns 2009).

1.2. Evaluación del estado del hierro

En el laboratorio se valoran una serie de parámetros bioquímicos y hematológicos que permiten estimar de forma correcta los niveles de hierro. Además, para completar el cuadro diagnóstico, esta información se debe tener en cuenta en el contexto de la historia nutricional y médica del paciente. Los parámetros generalmente utilizados para llevar a cabo esta evaluación son:

- La *Hemoglobina* es el principal transportador de oxígeno en sangre. Los niveles de hemoglobina son, por tanto, los más importantes a la hora de valorar la deficiencia de hierro y su último estadio, la anemia (WHO/CDC 2004, Wood y Ronnenberg 2006). La concentración de hemoglobina se expresa en gramos por 100 mL de sangre (g/dL). La distribución normal de hemoglobina varía con la edad, el sexo y el estado fisiológico (por ejemplo, durante el embarazo). Así, se han propuesto diferentes límites de normalidad, que van de 13,0 a 14,2 g/dL en hombres y de 11,6 a 12,3 g/dL en mujeres (WHO 2001, Beutler y Waalen 2006). Sin embargo, la hemoglobina por sí sola no se considera un factor adecuado para valorar el

estado del hierro, por su falta de sensibilidad a los cambios en los niveles de este elemento y porque no sólo la deficiencia de hierro causa anemia (Cook 2005).

- La *ferritina sérica* es un buen indicador de las reservas del hierro en el organismo. Por eso, la Organización Mundial de la Salud (OMS) emplea este valor y el de la hemoglobina para establecer los fenotipos clínicos de la anemia por deficiencia de hierro o anemia ferropénica (WHO 2011a, WHO 2011b). Aunque se han propuesto varios valores de corte, en esta tesis doctoral se ha considerado que un valor por debajo de los 20 µg/L representa un estado de deficiencia de hierro (Navas-Carretero et al. 2008, McLaren et al. 2011). En cualquier caso, la ferritina es una proteína reactante de fase aguda. Esto significa que sus niveles séricos pueden encontrarse elevados en caso de infección o inflamación, al margen de cualquier cambio que se pueda producir en las reservas reales de hierro.
- El *hematocrito* es el porcentaje (%) de sangre total compuesto por glóbulos rojos intactos o hematíes (Ruiz Argüelles y Ruiz Delgado 2014). Este parámetro se reduce en condiciones de deficiencia de hierro y anemia. Sin embargo no aporta información adicional que no pueda ser obtenida mediante la valoración de la concentración de hemoglobina (WHO 2008)
- El *hierro sérico* es la medida más obvia del estado del hierro en el organismo, los valores de hierro varían en función de diferentes factores incluyendo la concentración de su proteína de unión, la transferrina (Wells 2005). Existen, de forma general, tres modos de medir el hierro sérico: a) valorar el contenido total de hierro por unidad de volumen µg/dL; b) medir el número total de sitios de unión a la transferrina, conocido como capacidad de unión de hierro en µg/dL; c) estimar el porcentaje de los dos sitios de unión en todas las moléculas de transferrina que están ocupadas, lo que se denomina saturación de la transferrina (Abbaspour et al. 2014). Estas medidas pueden ser diferentes como resultado de variaciones circadianas, de la absorción de hierro diario, presencia de infección o procesos inflamatorios.

- El *Volumen Corpuscular Medio* (VCM) refleja el tamaño medio de los glóbulos rojos, expresado en femtolitros ($1 \text{ fL} = 10^{-15}$ litros). El VCM permite diferenciar las anemias causadas por deficiencias nutricionales, ya que disminuye en el caso de la anemia ferropénica (anemia microcítica) y aumenta en el caso de la anemia causada por deficiencia de vitamina B12 o ácido fólico (anemia macrocítica) (Ruiz Argüelles y Ruiz Delgado 2014).
- Otro buen indicador del estado del hierro es la concentración en plasma del *receptor de transferrina* (TfR). Este índice se relaciona con la intensidad de la eritropoyesis, ya que los receptores TfR se generan en el desarrollo de los glóbulos rojos (Abbaspour et al. 2014). En caso de deficiencia y reserva escasa de hierro, los niveles de TfR podrían aumentar, indicando una insuficiencia severa de hierro. Además, el TfR parece verse menos afectado por la inflamación que la ferritina sérica (Beguin 2003).

1.3. Trastornos del metabolismo del hierro

Como se mencionó en el apartado 1.1.3, el organismo goza de mecanismos de control altamente elaborados para mantener la homeostasis del hierro, regulando de forma coordinada la absorción, el reciclaje y la movilización del hierro almacenado. La alteración de cualquiera de estos procesos puede producir, bien anemia ferropénica cuando hay baja disponibilidad de hierro para sintetizar hemoglobina, o bien desórdenes de exceso de hierro relacionados con patologías en diferentes órganos.

1.3.1. Anemia ferropénica

La deficiencia funcional de hierro se asocia con baja disponibilidad de este mineral para la síntesis de hemoglobina. Esto acaba resultando en anemia, alterándose de forma perjudicial el suministro normal de oxígeno a los tejidos (Nairz y Weiss 2006). La anemia causada por deficiencia del hierro, conocida como anemia ferropénica, es un desorden pandémico. Según la OMS, un cuarto de la

población mundial padece anemia, siendo las mujeres en edad fértil el grupo más afectado. Se ha estimado que la prevalencia de hierro en mujeres jóvenes es del 8% al 33% en Europa, del 10% al 16% en EEUU y alrededor del 42% en los países en vías de desarrollo (WHO 2008). La anemia es un desorden clínico caracterizado por un descenso de los niveles de hemoglobina. En esta tesis los valores de hemoglobina (Hb, en g/dL) y ferritina (Ft, en $\mu\text{g/L}$), basados en la OMS y otros autores, con los que se definen los fenotipos clínicos de la anemia ferropénica son: normal ($\text{Hb} \geq 12$, $\text{Ft} \geq 20$), deficiente ($\text{Hb} \geq 12$ y $\text{Ft} < 20$ ó $\text{Hb} \leq 12$ y $\text{Ft} > 20$) y anémico ($\text{Hb} < 12$ y $\text{Ft} < 20$) (Navas-Carretero et al. 2008, McLaren et al. 2011, WHO 2011a, WHO 2011b).

La deficiencia de hierro no se encuentra siempre ligada a la anemia. Ciertos cambios funcionales pueden producirse sin llegar a desarrollar anemia, pero la mayoría de los síntomas clínicos como la fatiga, la debilidad, la falta de concentración, etc. aparecen con este desorden (Camaschella 2015). Los síntomas se manifiestan incluso en las formas leves y moderadas de anemia ferropénica, afectando negativamente funciones como el desarrollo cognitivo, la respuesta inmunitaria y la capacidad de trabajo. Por otro lado, la deficiencia de hierro durante el embarazo genera consecuencias adversas tanto en la madre como en el feto, pudiéndose aumentar el riesgo de infecciones, bajo peso al nacer, mortalidad materna y/o perinatal. La anemia ferropénica también causa reducción de la capacidad de aprendizaje y ratios más altos de morbilidad (Abbaspour et al. 2014). Por último, este desorden ha sido asociado a otras patologías como la osteoporosis (Toxqui y Vaquero 2015), la enfermedad de Alzheimer (Crespo et al. 2014), la aterosclerosis (Galesloot et al. 2015) y la trombosis (Oboshi et al. 2015).

1.3.2. Otros tipos de anemia

La presente tesis se centra en la anemia ferropénica, sin embargo, a continuación se describen sucintamente otros tipos de anemia:

Anemia por enfermedad crónica. Es el segundo tipo de anemia más frecuente (Zarychansky y Houston 2008). Los procesos inflamatorios prolongados en el

tiempo generan que la hepcidina se mantenga en valores altos, con lo que se bloquea el reciclaje en los macrófagos y la absorción de hierro (Weiss y Goodnough 2005).

Anemia por malnutrición. El organismo necesita vitaminas o minerales para generar eritrocitos. Para producir hemoglobina se requiere, además de hierro, vitamina B12, vitamina A, ácido fólico y cobre. Por tanto, la falta de cualquiera de estos elementos puede provocar una producción inadecuada de glóbulos rojos y, en consecuencia, anemia (Abbaspour et al. 2014).

Anemia falciforme. Esta enfermedad está causada por la producción de moléculas de hemoglobina con estructura anómala. Esto, a su vez, provoca una pérdida de afinidad de los glóbulos rojos por el oxígeno (Frenette y Atweh 2007).

Anemia por deficiencia refractaria al hierro. Es un trastorno autosómico recesivo poco frecuente, que se caracteriza por una anemia por deficiencia que no responde a la ingesta de hierro y con respuesta parcial al tratamiento parenteral de hierro (Orphanet 2015). Es causada por mutaciones en la matriptasa 2 que, al perder su función proteasa sobre la hemojuvelina (Figura 2), no evita que ésta mantenga los niveles altos de hepcidina (Du et al. 2008).

Anemia por hipotransferrinemia. Es un trastorno autosómico recesivo muy poco frecuente, debido a mutaciones en el gen que codifica para la TF, produciéndose una fuerte reducción de su síntesis (OMIM 2015). Esta reducción en los niveles de la TF conduce a una disminución del aporte de hierro en la médula ósea para el desarrollo de los precursores eritroides, lo que resulta en una síntesis de la hemoglobina reducida, y en la acumulación de un exceso de hierro en los tejidos periféricos. Los pacientes afectados presentan a la vez anemia hipocrómica microcítica y síntomas de hemocromatosis.

1.3.3. Hemocromatosis

La hemocromatosis es una patología que se caracteriza por los daños tisulares producidos por la sobrecarga de hierro. Puede ser primaria, como la

hemocromatosis genética, o secundaria cuando es el resultado de otra enfermedad asociada. Las enfermedades por sobrecarga de hierro suelen generarse de dos formas distintas. Una, cuando la eritropoyesis es normal y el hierro sérico excede la capacidad de unión a la transferrina, por lo que se deposita en las células parenquimatosas de los órganos. Otra, cuando la sobrecarga férrica es el resultado del aumento del catabolismo de los eritrocitos y el hierro se deposita inicialmente en las células del sistema mononuclear fagocítico y después en las células parenquimatosas.

La hemosiderosis se define como la sobrecarga de hierro que no genera daño tisular. Como se ha comentado anteriormente, el ión hierro es altamente reactivo y puede formar radicales libres que acaban por dañar los tejidos en los que se acumulan. El organismo humano no cuenta con un mecanismo activo de excreción de hierro, con lo que cualquier alteración en los niveles de absorción o en las proteínas que responden a diferentes niveles de hierro, podría resultar en una acumulación anormal de hierro en los órganos. El hígado es el órgano más afectado, debido a que es el tejido que más acumula este mineral, pero otros órganos también se pueden ver dañados. Un criterio general para considerar sobrecarga de hierro es presentar unos valores de índice de saturación de la transferrina mayor de 50%, junto con valores de ferritina sérica superiores a 150µg/L en mujeres y a 200µg/L en hombres (Bacon et al. 2011).

En función del tipo hemocromatosis, el momento de aparición y la gravedad de los síntomas varían. El cuadro clínico compartido por las diferentes hemocromatosis se caracteriza de forma generalizada por:

- Daños en el hígado que acarreen hepatomegalia y, en menor frecuencia, cirrosis. En enfermos cirróticos se desarrolla con frecuencia cáncer de hígado.
- Daños en el páncreas que pueden generar diabetes.
- Hipogonadismo secundario por daños en el eje hipotálamo-hipofisario.
- Afectación cardíaca por daños en el miocardio que afecta predominantemente a los ventrículos y al tejido contráctil. Esto deriva en insuficiencia cardíaca y arritmias.

- Artropatía que afecta fundamentalmente a las articulaciones metacarpofalángicas.
- Pigmentación cutánea por aumento de la melanina y de los depósitos de hierro subcutáneos (Bacon et al. 2011).

Existen diferentes tipos de hemocromatosis pero, dados los objetivos de la presente tesis doctoral, a continuación se procede a describir sólo las variantes de hemocromatosis hereditarias (HH):

Hemocromatosis hereditaria tipo I o debida a HFE: Este tipo de hemocromatosis es la más frecuente y su base genética se conoce desde hace varias décadas (Feder et al. 1996). Se origina debido a mutaciones del gen *HFE*, en concreto C282Y y H63D. Las diferentes combinaciones genotípicas son responsables de la gravedad y penetrancia de la enfermedad. Alrededor del 80-90% de individuos caucásicos con hemocromatosis hereditaria presentan C282Y en homocigosis. Los dobles heterocigotos C282Y/H63D presentan también altas probabilidades de desarrollar la enfermedad (Feder et al. 1996, Pietrangelo 2004). Sin embargo, la penetrancia de la mutación C282Y en homocigosis parece ser alrededor del 24 al 43%, además de ser una enfermedad más frecuente en hombres. Es la penetrancia de la sobrecarga la que sí parece ser cercana al 100%, lo que evidencia que hay otros muchos factores genéticos y ambientales que influyen en su desarrollo (Rossi et al. 2008). El gen *HFE* codifica una proteína transmembrana no clásica de la familia del complejo de histocompatibilidad que interacciona con TfR1 y con el TfR2 como se vio en el apartado 1.1.3. La mutación C282Y causa una sustitución de una cisteína por una tirosina y genera una modificación en la estructura terciaria que hace perder la funcionalidad de la proteína (Figura 4). Así, deja de interactuar con los receptores de transferrina y, por tanto, cesa su capacidad de estimular la expresión de hepcidina, de forma que el hierro entra de forma descontrolada al interior celular (Ortiz et al. 2012).

Hemocromatosis hereditaria tipo II o Hemocromatosis juvenil: Es considerada una enfermedad rara, como todas las demás HH salvo la tipo I. Se caracteriza por una aparición temprana de los síntomas de la enfermedad (cirrosis, insuficiencia

cardíaca, infantilismo etc.), debido a una excesiva absorción intestinal de hierro (Cazzola et al. 1998). Afecta tanto a varones como a mujeres, antes de los 30 años y es predominante en sujetos de ascendencia europea. La base genética de la enfermedad se caracteriza por mutaciones en el gen de la hemojuvelina (Tipo IIa) y hepcidina (Tipo IIb). El primer estudio en identificar una alteración en la hemojuvelina fue el de Papanikolau et al. (2004). El gen de la proteína *HFE2*, como este autor inicialmente denominó a la hemojuvelina, presentaba la mutación G230V en varios individuos de diferente origen. Desde entonces, se han detectado varias mutaciones, tanto en hemojuvelina como en hepcidina, ya sea en regiones codificantes o intergénicas (Roetto et al. 2003, Huang et al. 2004, Lee et al. 2004, Lanzara et al. 2004, Murugan et al. 2008, Brakensiek et al. 2009). Todas estas mutaciones están restringidas a ciertas familias y su frecuencia a nivel poblacional es muy baja.

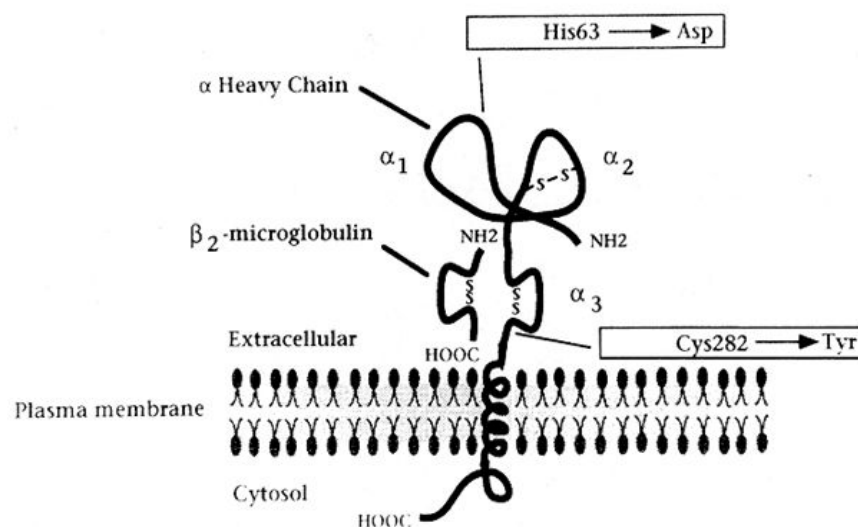


Figura 4. Proteína HFE y sus principales mutaciones. La mutación C282Y, la más grave de la hemocromatosis hereditaria tipo I, genera la ruptura de un puente de disulfuro que compromete su estructura terciaria y, por tanto, su función. Tomado de Bacon et. al 2011

Hemocromatosis hereditaria tipo III asociada al receptor 2 de la transferrina: Con sintomatología similar a la HH tipo II, se debe a mutaciones en el gen que codifica para el receptor 2 de la transferrina. Inicialmente descubierta en familias

procedentes del sur de Italia (Camaschella et al. 2000, Pietrangelo 2005, Radio et al. 2014), también se ha descrito en Japón (Hattori et al. 2003), Portugal (Mattman et al. 2002) y norte de Francia (Le Gac y Férec 2005).

Hemocromatosis hereditaria tipo IV enfermedad de la ferroportina: Al contrario del resto de las HH, ésta presenta una modelo de herencia autosómica dominante. Los niveles de ferritina suelen estar muy elevados mientras que la transferrina se puede mantener normal (Devalia et al. 2002, Roetto et al. 2002). Esta disparidad parece ser más marcada en fases tempranas de la enfermedad. El análisis del gen de la ferroportina reveló varias mutaciones causantes de esta variante de la HH (Rivard et al. 2003).

1.4. Marcadores genéticos relacionados con el metabolismo del hierro

1.4.1. Heredabilidad y desequilibrio de ligamiento

La heredabilidad es un concepto matemático asociado a la varianza estadística. Nombrada como H^2 , es la proporción de varianza total de un determinado parámetro, explicada por la varianza genética en una población dada (Ramírez y Egaña 2003). Se ha empleado, de forma un tanto imprecisa, como indicador del grado de control genético al que está sometido un determinado fenotipo. Esto no es necesariamente correcto porque hay fenotipos poco variables, con heredabilidad baja, cuya expresión depende casi exclusivamente de factores genéticos. Su utilidad radica en que, para aquellos factores con alta heredabilidad, será más probable encontrar marcadores genéticos relacionados con su variación a nivel poblacional. Éste el motivo por el cual en los estudios de asociación se hace una valoración previa de heredabilidad de los fenotipos que se quieren estudiar.

El desequilibrio de ligamiento (LD, *linkage disequilibrium*) es un parámetro que describe el patrón de asociación no aleatoria, entre alelos de diferentes loci (regiones cromosómicas) en una determinada población (Nordborg y Tavaré 2002). Se podría decir que es un concepto con una definición confusa, tal y como

otros autores han mostrado, ya que no está necesariamente relacionado con el ligamiento biológico, ni tampoco con la falta de equilibrio (Slatkin 2008). Términos más adecuados para definir este fenómeno serían la “asociación alélica no aleatoria” o “fase gamética”. Dadas unas frecuencias alélicas, cuando dos alelos en diferentes loci se presentan en una frecuencia conjunta más alta de la esperada en una determinada población, se dice que esos dos loci están en LD. Esta frecuencia conjunta esperada estaría definida por el producto de las frecuencias alélicas de cada marcador. Es, por tanto, un parámetro estrictamente poblacional, mientras que el ligamiento biológico es un fenómeno intrínseco al ADN. Este término se usó por primera vez en 1960 por Lewontin y Kojima, y persiste porque su uso se extendió en el campo de la genética de poblaciones, donde el término se dio por válido mientras la definición estadística quedara clara. Sin embargo, en los años 80 el interés por el LD creció rápidamente dado su utilidad en los estudios de genotipado y, para entonces, el término ya estaba demasiado establecido para ser cambiado. Estadísticamente hay varias definiciones pero todas están basadas en la siguiente fórmula donde el parámetro D sería el indicador básico de LD:

$$D = P_{AB} - P_A P_B$$

Que representa la diferencia entre la frecuencia de los gametos que llevan el alelo A y B (P_{AB}) y el producto de las frecuencias de estos alelos ($P_A P_B$).

Además de la recombinación genética, dos alelos pueden estar asociados por múltiples causas que tienen que ver con la historia evolutiva de una determinada población y los eventos que influyen en ella, como por ejemplo la selección natural, la deriva génica, el efecto fundador y la estratificación poblacional, entre otras.

Desde el punto de vista de la genética clínica, el LD se ha revelado como fundamental en los estudios de asociación fenotipo-enfermedad, especialmente cuando se emplean técnicas de genotipado masivo (GWAS).

1.4.2. Variabilidad poblacional y proyecto *HapMap*

El proyecto *HapMap* se desarrolló a partir de los datos que se iban revelando de los patrones de bloques de haplotipos en desequilibrio de ligamiento en el genoma. En el año 2001, Jeffreys et al. y Daly et al., de manera independiente, mostraron como el genoma estaba estructurado en bloques de haplotipos, del orden de decenas o centenas de kilobases, dentro de los cuales la variabilidad era muy baja. En la Figura 5 se muestra que un haplotipo es un conjunto de marcadores localizados en fase o en el mismo cromosoma. Los tag-SNP (*Single Nucleotide Polymorphism*) o SNPs etiqueta son polimorfismos de un único nucleótido que recogen la mayoría de la variabilidad haplotípica de una determinada región. Además, estos bloques parecían estar flanqueados por puntos

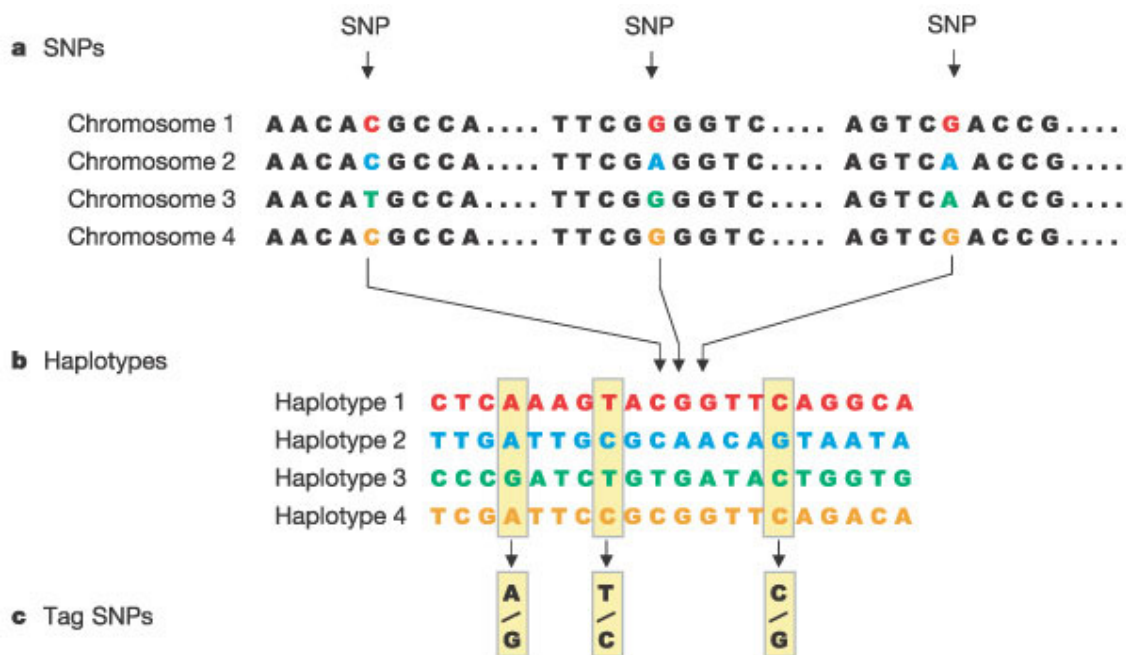


Figura 5. Descripción de la variabilidad en el genoma. a) Los polimorfismos de un solo nucleótido (SNPs) son la forma más común de variación en el ADN. b) Los SNPs forman haplotipos, de forma que la variación total de una región puede ser descrita mediante un número reducido de haplotipos. c) Los tag-SNPs o SNPs etiqueta son marcadores que permiten identificar la mayoría de los haplotipos incluidos en una región. Tomado de *International HapMap Consortium 2003*

de recombinación. Daly et al. (2001) diseñaron un sistema para valorar el LD en función de los bloques de haplotipos. El hecho de que esa forma de organizarse el genoma parecía generalizada, constituía un marco adecuado para identificar los haplotipos que definen todo el genoma. De esta forma, se reducía considerablemente la variación inicialmente esperada a tenor del gran número de SNPs contenidos en el ADN. Tal y como el mismo Consorcio Internacional de *HapMap* definió, el objetivo del proyecto *HapMap* (mapa de haplotipos) era determinar los patrones comunes de variación de la secuencia de ADN y poner toda esa información disponible para la comunidad científica de forma libre y gratuita (*International HapMap Consortium* 2003).

Este consorcio comunicó en 2003 la intención de secuenciar más de un millón variantes o marcadores del genoma en individuos de ascendencia africana, asiática y europea, publicar sus frecuencias y el grado de asociación entre estos marcadores. Como futuros avances esperados, el mapa de haplotipos (*HapMap*) contribuiría a encontrar variantes asociadas a enfermedades comunes, facilitaría el desarrollo de herramientas de diagnóstico y potenciaría la capacidad de elegir dianas para la intervención terapéutica. En 2005, el consorcio publicó el primer mapa de haplotipos obtenido a partir de 265 muestras. La información que se hacía disponible describía puntos calientes de recombinación, estructuras de bloques de LD y una baja diversidad de haplotipos que dejaba en evidencia las numerosas correlaciones que existían entre conjuntos de SNPs. Se mostraba también cómo esta información podía ayudar a diseñar estudios de asociación y a identificar marcadores objeto de selección natural (*International HapMap Consortium* 2007).

En la actualidad, la plataforma de *HapMap* ofrece información de más de 3,1 millones de SNPs, genotipados en 1301 individuos seleccionados a nivel mundial a partir de 11 poblaciones distintas. Con esta cobertura de genotipado, se estima que está cubierto de un 25 a un 35% de toda la variación de SNPs comunes en las poblaciones empleadas. La comentada plataforma de búsqueda online, aparte de ofrecer información sobre frecuencias genotípicas y alélicas, pone a disposición del usuario varias herramientas. Se menciona, por ejemplo, su utilidad para seleccionar tag-SNPs o conocer el patrón de desequilibrio de ligamiento de una

región de interés, extraer información sobre la fase haplotípica, estimar el grado de desequilibrio de ligamiento entre dos SNPs, etc. Además, se ofrecen enlaces a otras bases de datos y plataformas de interés, como el catálogo NHGRI de estudios de asociación para marcadores relacionados con enfermedades, la base de datos OMIM de enfermedades con herencia mendeliana, la plataforma *Entrez genes*, y otras (*International HapMap Consortium* 2007, plataforma online del proyecto *HapMap*: <http://hapmap.ncbi.nlm.nih.gov/>).

1.4.3. Estudios de asociación y base genética de las alteraciones del metabolismo del hierro

Para conocer el componente genético de las enfermedades, hasta hace pocos años se han empleado diferentes abordajes para estudiar rasgos mendelianos y monogénicos. La estrategia genes-candidato y los estudios de ligamiento familiar, en las que se emplean STRs y SNPs, obtenían buenos resultados para rasgos sencillos, pero en general se encontraban variantes poco comunes asociadas a patologías graves y raras (Carracedo y Salas 2008). En concreto, los estudios genes-candidato se basan en hipótesis previas sobre el conocimiento de que ciertos genes se relacionan con un determinado fenotipo. Sin embargo, esta estrategia es poco útil en el estudio de las enfermedades de herencia compleja o multigénica.

Los estudios de asociación mediante genotipado masivo, en inglés “*Genome-Wide Association Studies*” (GWAS), abrieron un nuevo campo en la investigación de la base genética de las enfermedades comunes de herencia compleja (Smith y Newton Cheh 2009). Gracias a los GWAS, se han podido descubrir señales de asociación en regiones inesperadas con rasgos de herencia compleja o poligénica, que más tarde han podido ser replicadas (Carracedo y Salas 2008). Esta técnica se basa en realizar un genotipado masivo de variantes polimórficas (comunes) en individuos que presenten diferentes fenotipos y, posteriormente, tratar de asociar esas variantes con un determinado fenotipo. Debido a que en nuestra historia evolutiva el primer ancestro común es relativamente reciente, y que los eventos de

recombinación se suelen concentrar en determinadas regiones cromosómicas, somos capaces de sondear toda la variabilidad común del genoma usando unos cuantos miles de marcadores o variantes. El patrón de desequilibrio de ligamiento es lo que nos permite usar los marcadores analizados para desvelar la asociación de marcadores no genotipados.

Por otro lado, estos estudios GWAS adolecen de ciertos problemas a la hora de replicar los resultados, hecho que guarda una estrecha relación con el diseño experimental y con ciertos aspectos estadísticos (Chanock et al. 2007). Estos problemas se manifiestan en errores tipo I, o falsos positivos, y en errores tipo II, que representan los falsos negativos. La *estratificación poblacional*, o estructuración de la muestra, puede acarrear asociaciones falsas (errores tipo I), debido a que los casos y controles son poblaciones distintas a nivel genético (Salas y Carracedo 2007). Las *correcciones por comparaciones múltiples* son necesarias, especialmente en el caso de analizar muchos SNPs, ya que las pruebas estadísticas se multiplican exponencialmente pudiéndose generar errores de tipo I. Sin embargo, una estrategia de corrección muy conservadora podría provocar una reducción en la potencia del estudio y, por ende, errores de tipo II. La *potencia* de la prueba o el *poder estadístico* de un estudio de asociación también viene determinada por el tamaño de la muestra. Una muestra no suficientemente grande podría ser la causa del “efecto *Jackpot*” relacionado con la falta de potencia estadística. Este efecto consiste en que el primer estudio que informa de una asociación débil posiblemente haya sobrestimado el efecto real del polimorfismo, con lo que otros estudios fracasan a la hora de replicarlo. Una correcta *definición del fenotipo* de interés también puede ser fuente de errores estadísticos. El hecho de que bajo una misma definición convivan varias alteraciones con base genética distinta conduce a una pérdida de detección (Carracedo y Salas 2008).

En el marco de los estudios de asociación, también se encuentran los análisis de “mapeo fino” o *Fine Mapping*. De forma general, un *Fine Mapping* es un análisis genético, en el que se genotipan varios SNPs de una región que anteriormente haya resultado asociada con algún fenotipo de interés. Lo que se busca es profundizar en el conocimiento de esa asociación, tratando de encontrar el SNP responsable o causal del hallazgo estadístico. La selección de los SNPs se realiza teniendo en

cuenta el LD de la región, de manera que se pueda captar la máxima variación, genotipando el menor número de SNPs posibles.

En cuanto al componente genético de las alteraciones del estado del hierro, se han descrito numerosas mutaciones en proteínas implicadas en el metabolismo, en la estabilidad del eritrocito o en la absorción del hierro que producen anemia severa o hemocromatosis (Weatherall 1998, Roetto et al. 2002, Gehrke et al. 2005, Finberg 2009, Ramsay et al. 2009, de Tayrac et al. 2015). La heredabilidad de los parámetros empleados para valorar el estado del hierro (Hb, Ft, Tf) son reflejo de un componente genético, que podría explicar parte de la variabilidad de estos parámetros y por tanto del estado del hierro (Whitfield et al. 1985, Garner et al. 2000, Souto et al. 2003, Njajou et al. 2006, Sala et al. 2008). En última instancia, esto ha hecho posible encontrar en los estudios de asociación variantes comunes asociadas con las alteraciones del estado del hierro (Bayele et al. 2009). Trabajos como el de Benyamin et al. (2009a) revelaron cómo un número discreto de marcadores en el gen de la transferrina era capaz de explicar una proporción muy elevada de la variación genética total de esta proteína. Otros trabajos de asociación, empleando grupos familiares como muestra, señalan cómo dos regiones mostraban señales de ligamiento con la variación de los niveles del TfR1: la más alta en cromosoma 2 y otra ligeramente más baja en cromosoma 3 (Remacha et al. 2006).

Los estudios de genes candidato en enfermedades de herencia monogénica también revelaron mutaciones que se localizaban en regiones codificantes de los genes relacionados con el metabolismo del hierro (Carracedo y Salas 2008). A partir del empleo extendido de los GWAS, se han identificado numerosas variantes comunes (SNPs) localizadas en regiones intergénicas que regulan el estado del hierro, identificándose, más que mutaciones causales, factores de riesgo y protección para ciertas alteraciones del hierro. La abundancia además de la relevancia de estas variantes intergénicas e intrónicas se vio reforzada con los resultados del proyecto *Encode* (Dunham et al. 2012, Maurano et al. 2012). Además, debido al diseño de los GWAS sin hipótesis previa, se han revelado regiones y genes que, a priori, nunca hubieran sido estudiadas en relación al estado

del hierro, profundizándose así, el conocimiento sobre las bases de este metabolismo.

A modo de revisión somera, se han hallado numerosas asociaciones significativas entre SNPs y los parámetros hematológicos y bioquímicos empleados para valorar el estado del hierro como la hemoglobina, ferritina, hematocrito, volumen corpuscular medio, transferrina, y otros. Estos marcadores están repartidos por todo el genoma, aunque muchos de ellos se han encontrado en el cromosoma 6 (Benyamin et al. 2009b, Soranzo et al. 2009, Ganesh et al. 2009, Chambers et al. 2009, Kullo et al. 2010, Blanco-Rojo et al. 2011, Lo et al. 2011, Oexle et al. 2011, Pichler et al. 2011, McLaren et al. 2012, Van der Harst et al. 2012, Benyamin et al. 2014, McLaren et al. 2015, Li et al. 2015). Además, aunque estos estudios son más escasos, también se han encontrado asociaciones de marcadores con el riesgo de padecer anemia ferropénica (McLaren et al. 2011, An et al. 2012, Blanco-Rojo et al. 2014).

1.5. Objetivos

El objetivo general de esta tesis doctoral es profundizar en el conocimiento de la asociación entre marcadores genéticos y los trastornos del metabolismo del hierro.

OBJETIVOS ESPECÍFICOS:

1. Diseñar y valorar una herramienta de genotipado (reacción de mini-secuenciación *SNaPshot*), para replicar la asociación entre marcadores genéticos y parámetros bioquímicos relacionados con el metabolismo del hierro.
2. Comparar la variabilidad de los marcadores incluidos en el *SNaPshot* entre población española y otras caucásicas. Estudiar su posible estructura poblacional, con el fin de valorar la utilidad clínica de estas variantes.
3. Buscar variantes en los genes candidato del exceso de hierro y/o hemocromatosis (hepcidina y hemojuvelina) en individuos con diferentes genotipos para el gen *HFE*. Comparar su distribución alélica con la de otras poblaciones.
4. Establecer asociaciones entre marcadores genéticos y parámetros bioquímicos y hematológicos relacionados con el metabolismo del hierro.
5. Estudiar las asociaciones obtenidas más relevantes, con el fin de proponer nuevas hipótesis sobre su implicación en la etiología de la anemia ferropénica.
6. Establecer un modelo para identificar factores genéticos de riesgo o protección que influyen en el desarrollo de la anemia ferropénica.

2. DESARROLLO DE UNA HERRAMIENTA DE GENOTIPADO PARA DETECTAR MARCADORES GENÉTICOS RELACIONADOS CON EL METABOLISMO DEL HIERRO

2.1. “A novel SNaPshot assay to detect genetic mutations related to iron metabolism”

2011

Genet Test Mol Biomarkers. 2011 Mar;15(3):173-9

doi: 10.1089/gtmb.2010.0140

Epub 2011 Jan 3

PMID: 21198378

GENETIC TESTING AND MOLECULAR BIOMARKERS
Volume 15, Number 3, 2011
© Mary Ann Liebert, Inc.
Pp. 173–179
DOI: 10.1089/gtmb.2010.0140

A Novel SNaPshot Assay to Detect Genetic Mutations Related to Iron Metabolism

Stefania Bertoni¹, Ruth Blanco-Rojó,² Carlos Baeza,³ Eduardo Arroyo-Pardo,³ María Pilar Vaquero,² and Ana María López-Parra³

The minisequencing method is a cost-effective tool to study single-nucleotide polymorphisms in human disease. For this reason, a novel polymerase chain reaction multiplex SNaPshot reaction has been developed that targets 10 autosomal mutations in genes, or regions near to them, reported to be involved in iron metabolism: TMPRSS6, TF, and HFE. To validate this multiplex, 284 samples from unrelated women from the Spanish population were tested at a fertile age. The novel polymerase chain reaction multiplex SNaPshot reaction developed is a very simple, sensitive, and low-cost approach, and therefore will be useful as a tool of clinic diagnosis in iron metabolism alterations and to replicate results obtained from genome-wide linkage analysis.

Introduction

IRON IS AN essential element for biochemical functions and its overload can cause liver diseases, as hemochromatosis, whereas iron deficiency can lead to anemia. Iron deficiency and iron-overload are widespread health problems that affect the quality of life with significant social and economic impact. Both of them are multifactorial diseases in which, among others, genetic, environment, and nutritional factors are involved. Iron status is assessed by measuring hematological and biochemical parameters such as hemoglobin, serum iron, serum transferrin, transferrin saturation (SAT), soluble transferrin receptor (sTfR), and serum ferritin. Remacha *et al.* (2006) indicated that heritability of iron phenotypes ranged from 0.18 in the case of SAT to 0.55 in sTfR using family-based studies. Therefore, 18%–55% of the total variability of iron parameters was controlled by genes.

Single-nucleotide polymorphisms (SNPs) are the most common type of genetic variation and they are widely represented in the genome, but only a large portion of them are genuine polymorphisms. About 3 million SNPs from human population samples have been identified, and it is believed that the human genome contains at least 10 million SNPs. Some of these variants could be genetic factors that lead to iron overload or deficiency. High-throughput techniques have been recently developed to detect human mutations related to human disease. Among common applications, there are the association studies between bi-allelic variants of SNPs and differential susceptibility to develop certain diseases or

(Bustamante-Aragones *et al.*, 2008; Barden *et al.*, 2009). Another application of great interest is the identification of haplotype blocks in the human genome to simplify genetic analysis for use in disease diagnosis and human genetic variation (Patil *et al.*, 2001; Stephens *et al.*, 2001; Gabriel *et al.*, 2002).

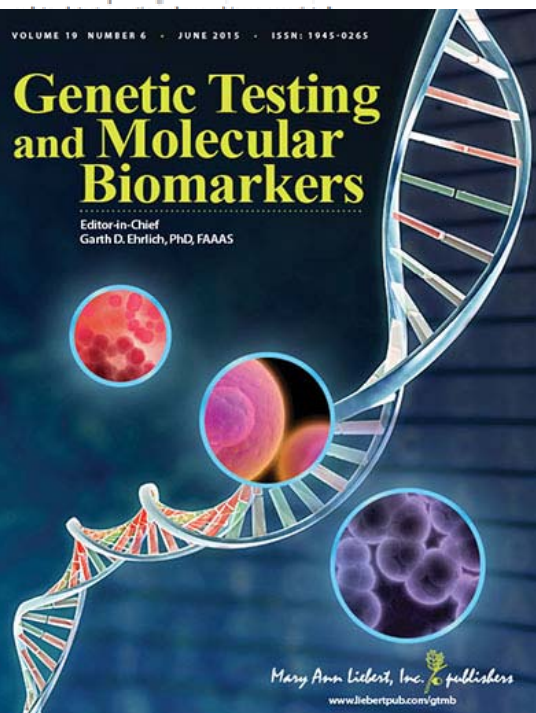
Previous studies of our research group discussed the approach to single mutations of iron metabolism genes in association studies and the possible gene-gene interactions (Sarría *et al.*, 2007a; Navas-Carretero *et al.*, 2008). Genome-wide association studies (GWASs) and genome-wide linkage analysis permit identification of genetic associations with observable traits (quantitative trait locus). In most of these studies, only a small number of SNPs are associated with a disease or a trait. Recently, some GWASs for serum markers of iron status, that is, serum iron, serum transferrin, SAT, sTfR, and serum ferritin, have been published (Remacha *et al.*, 2006; Benyamin *et al.*, 2009; Chambers *et al.*, 2009; Ganesh *et al.*, 2009; Soranzo *et al.*, 2009). These results need to be replicated in other samples for which simple and inexpensive techniques may be used. The minisequencing method (SNaPshot kit) is

currently with snp
2009). T
tively si
The p
to be us
found in
metabol
a backb

VOLUME 15 NUMBER 6 • JUNE 2011 • ISSN: 1045-0265

Genetic Testing and Molecular Biomarkers

Editor-in-Chief
Garth D. Ehrlich, PhD, FAAAS



Mary Ann Liebert, Inc. publishers
www.liebertpub.com/gtmb

2.1 "A novel SNaPshot assay to detect genetic mutations related to iron metabolism" Artículo

Bertoncini Stefania¹, Blanco-Rojo Ruth², Baeza Carlos³, Arroyo-Pardo Eduardo³, Vaquero M. Pilar², López-Parra Ana Maria³

¹ Department of Biology, Anthropology Unit, University of Pisa, Pisa, Italy.

² Department of Metabolism and Nutrition, ICTAN, Spanish National Research Council (CSIC), Madrid, Spain.

³ Department of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Spain.

Author for correspondence: Ana María López Parra, e-mail: amlopezparra@med.ucm.es

Resumen

La minisequenciación es un método económico y efectivo para estudiar polimorfismos de un sólo nucleótido involucrados en el desarrollo de determinadas patologías. Por ello se ha desarrollado un *multiplex SNaPshot* para genotipar de forma simultánea 10 SNPs relacionados con el metabolismo del hierro y localizados en genes como la matriptasa, la transferrina y el *HFE*. Para validar el método se genotiparon 284 mujeres en edad fértil de población española. Esta nueva herramienta *SNaPshot multiplex* es sencilla de implementar, económica, y de gran sensibilidad. Por tanto, será de utilidad para el diagnóstico clínico de las alteraciones del metabolismo del hierro y para replicar los resultados de los estudios de asociación de genotipado masivo.

Palabras clave: *polimorfismo de un solo nucleótido, metabolismo del hierro, multiplex PCR, SNaPshot, población española, extensión de base simple*

Abstract

The minisequencing method is a cost-effective tool to study single nucleotide polymorphisms (SNPs) in human disease. For this reason, a novel PCR multiplex SNaPshot reaction has been developed that targets 10 autosomal mutations in

genes, or regions near to them, reported to be involved in iron metabolism; TMPRSS6, TF and HFE. To validate this multiplex, 284 samples from unrelated females from the Spanish population, were tested at a fertile age. The novel PCR multiplex SNaPshot reaction developed is a very simple, sensitive and low cost approach, and therefore will be useful as a tool of clinic diagnosis in iron metabolism alterations and to replicate results obtained from genome wide linkage analysis.

Keywords: *single nucleotide polymorphisms, iron metabolism, multiplex PCR, SNaPshot, Spanish population, single base extension*

Introduction

Iron is an essential element for biochemical functions and its overload can cause liver diseases, as hemochromatosis, whereas iron deficiency can lead to anemia. Iron deficiency and iron-overload, are widespread health problems that affect the quality of life with significant social and economic impact. Both of them are multifactorial diseases in which among others, genetic, environment and nutritional factors are involved. Iron status is assessed by measuring haematological and biochemical parameters such as: haemoglobin, serum iron, serum transferrin, transferrin saturation (SAT), soluble transferrin receptor (sTfR), and serum ferritin. Remacha *et al.* (2006) indicated that heritability of iron phenotypes ranged from 0.18 in the case of SAT to 0.55 in sTfR using family-based studies. Therefore, 18–55% of the total variability of iron parameters was controlled by genes.

Single nucleotide polymorphisms (SNPs) are the most common type of genetic variation and they are widely represented in the genome, but only a large portion of them are genuine polymorphisms. About 3 million SNPs from human population samples have been identified and it is believed that the human genome contains at least 10 million SNPs. Some of these variants could be genetic factors that lead to iron overload or deficiency. High-throughput techniques have been

recently developed to detect human mutations related to human disease. Among common applications, there are the association studies between bi-allelic variants of SNPs and differential susceptibility to develop certain diseases or responsiveness to drug therapy in different human groups (Bustamante-Aragones *et al.*, 2008; Bardien *et al.*, 2009). Another application of great interest is the identification of haplotype blocks in the human genome to simplify genetic analysis for use in disease diagnosis and human genetic variation (Patil *et al.*, 2001, Stephens *et al.*, 2001, Gabriel *et al.*, 2002).

Previous studies of our research group discussed the approach to single mutations of iron metabolism genes in association studies and the possible gene-gene interactions (Sarriá *et al.*, 2007a; Navas-Carretero *et al.*, 2008). Genome wide association studies (GWASs) and genome wide linkage analysis permit identification of genetic associations with observable traits (quantitative trait locus or QTL). In most of these studies, only a small number of SNPs are associated with a disease or a trait. Recently some GWASs for serum markers of iron status; i.e., serum iron, serum transferrin, SAT, sTfR, and serum ferritin have been published (Remacha *et al.*, 2006; Benyamin *et al.*, 2009; Chambers *et al.*, 2009; Ganesh *et al.*, 2009, Soranzo *et al.*, 2009). These results need to be replicated in other samples for which simple and inexpensive techniques may be used. The minisequencing method (SNaPshot kit) is currently used to detect genetic polymorphisms associated with small sets of SNPs (Filippini *et al.*, 2007, Ferri and Pelotti, 2009). The advantages of this technique are mainly its relatively simple optimization, low cost and rapid results.

The purpose of this work was to design a multiplex system to be used to replicate the association signal of some SNPs, found in GWAS, with biochemical parameters related to iron metabolism. We also attempt to validate the system by typing a healthy unrelated population of females.

Subject and methods

SNPs selection

A set of 10 SNPs was selected (rs4820268, rs855791, rs1799852, rs2280673, rs1800562, rs3811647, rs2673289, rs1375515, rs1799945, rs16826756). Six of these SNPs were chosen due to their association with parameters related to iron metabolism, from the GWAS of Benyamin *et al.* (2009) (rs3811647, rs1799852, rs2280673, rs4820268, rs1800562, rs855791). The HFE mutation H63D (rs1799945), which was not genotyped in the above mentioned GWAS, was included in our study due its important role in iron overload diseases (Feder *et al.*, 1996). The last three SNPs (rs2673289, rs1375515, rs16826756) were selected from two regions, which showed a linkage signal with the levels of sTfR, detected by Remacha *et al.* (2006). On-line software at HapMap Project (The HapMap Project: <http://www.hapmap.org/>) allowed selection of tag-SNPs that captured most of the variability of the two regions of interest. The features of these SNPs were obtained from the public database dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/>).

Primers design

Primers for PCR amplification were designed to have an annealing temperature between 55 and 60°C using Primer3 (<http://www.genome.wi.mit.edu/cgi-bin/primer/primer3>) and Aspe Tools software (<http://yellow.nist.gov:8444/dnaAnalysis/aspeToolsPage.do>). In order to test for possible nonspecific amplications, primers were aligned with the NCBI sequence databases using BLAST (Basic Local Alignment Search Tool): <http://www.ncbi.nlm.nih.gov/blast/blast.cgi>). AutoDimer Software (<http://www.cstl.nist.gov/div831/strbase//AutoDimerHomepage/AutoDimerProgramHomepage.htm>) was used to test for potential hairpin and primer-dimer structures. Detailed information about primers, probes, amplified products and the SNPs contain in each amplicon, can be seen in Table 1.

Multiplex PCR reaction

The primers were combined into a single multiplex reaction, generating different PCR product sizes (112-495bp). The best conditions obtained are provided in Table 1. PCR reaction was performed in a volume of 5µl using 10-20ng of genomic DNA, 2.5 µl of PCR Master Mix from QIAGEN Multiplex PCR kit (Qiagen Hilden Germany) and 0.3µl of 10x premixed primers. This Master Mix contains a unique PCR buffer with a novel synthetic factor that stabilizes specifically bound primers and enables efficient extension of all primers in the reaction without demanding optimization. The amplification consisted of 95°C for 15 min, followed by 31 cycles for 30 s at 94°C, 90 s at 58°C, 90 s at 72°C, and a final extension for 10 min at 72°C. The PCR products were resolved in 2.5% NuSieve GTG agarose gel.

Minisequencing SNaPshot reaction

In order to eliminate the excess of primers and dNTPs, 0.6µl of PCR product was purified with 0.9 µl of Exo-SAP-IT (U.S.B. Corporation Cleveland, OH, USA) and incubated at 37°C for 15 min. The enzymes were afterwards inactivated by heating at 80°C for 15 min.

In the minisequencing reaction, the SNaPshot primers were designed employing the same software used for the multiplex PCR. They present a poli-C or a poli-GACT tail at the 5' end, except for the first one. Size of adjacent single base extension primers varied between 5 and 6 bases to avoid overlapping caused by differences in electrophoretic mobility (Table 1). Multiplex PCR minisequencing was performed in a 5µl volume using 1.5 µl of purified PCR product, 0.04-0.78 µM of primers and 1µl of SNaPshot ready reaction mix. The PCR program consisted of 25 cycles for 10s at 96°C, 5s at 50°C and 30s at 60°C.

To eliminate unincorporated ddNTPs, each sample was purified by 1µl (0.2 Units/µl) of SAP enzyme (Shrimp Alkaline Phosphatase) and incubated at 37°C for 60 min and at 85°C for 15 min for inactivation.

Table. Mean Features of 10 Single-Nucleotide Polymorphisms, Polymerase Chain Reaction Amplification Reactions and Multiplex Single-Base Extension Reactions For The Single Polymerase Chain Reaction-Multiplex developed in this study

Name	SNP Primer Sequences 5'-3'	Length	Size	SNP Code	PCR concentrations (μM)	SBE Primer sequences 5'-3'	SBE Primer Length	Tail	Final length	Minisequencing SBE Primer Concentration (μM)
4820268f	aggagattgatggtgac	20	394	A/G	3	Rv-TTGTCTCAACGGCAGCGA	18		19	0,18
4820268r	cacagttgcagagccacat	20	394		3					
855791f	atgtgcagttgatccacag	20	112	C/T	1	Rv-CGTCACTGGTAGCGATAG	19	(dC)5	25	0,1
855791r	ctcactgcagggcctcctt	20	112		1					
1799852f	ttccctcccagagaactt	20	312	C/T	1	Fw-ACCAGTATGAGCTGCTTTGC	20	(dC)9	30	0,26
1799852r	catggctcccactgtatcct	20	312		1					
2280673f	gcacattggtgcagacactt	20	364	A/C	1	Rv-GGTATCTCATCTGTGAAGCCC	21	(dC)14	36	0,16
2280673r	ccaccctccactcctctgta	20	364		1					
1800562f	tacccccagaaacatcacat	20	225	A/G	1	Fw-GGGGAAGAGCAGAGATATACGT	22	(dC)19	42	0,04
1800562r	cctggctctcagtcacaca	20	225		1					
3811647f	ggcttgacacaggattttt	20	495	A/G	2	Fw-CTGAGGGAGTTTACAGACAGATC	23	(dC)24	48	0,32
3811647r	agtgaagtggaaggacag	20	495		2					
2673289f	cattaataggctggcgatgg	20	486	C/T	8	Fw AGC/AGATCCTGTCTCAG/ATGTA	24	(dC)28	53	0,78
2673289r	gtgctgccacagaaacagac	20	486		8					
1375515f	ccttcgctgttcacata	20	347	A/G	1	Fw-CTACATGAATATCAGCCCAATTGAC	24	(dC)33	58	0,4
1375515r	ctaaagatggggcttggga	20	347		1					
1799945f	acatggttaaggcctgttgc	20	245	C/G	0,8	Fw-CCAGCTGTTCGTTCCTATGAT	22	(dC)41	64	0,4
1799945r	atgtagtccaccctttcag	20	245		0,8					
16826756f	ttcatcttatactgacctgagc	23	152	A/G	3	Fw-TTCAAAATTAGGTCCAAATTGCTACATAA	27	(dGACT)11	72	0,4
16826756r	gtctggccagcactttcaac	20	152		3					

PCR, polymerase chain reaction; SNP, single nucleotide polymorphism

Capillary electrophoresis and data analysis

The minisequencing purified products (0.5 µl) were mixed with 9µl of HiDi™ formamide and 0.5 µl of GeneScan-120 LIZ size standard (Applied Biosystems) and denatured at 95°C for 5 minutes. The fluorescently labelled fragments were resolved by capillary electrophoresis on an ABI PRISM 310 DNA Genetic Analyser (Applied Biosystems, Foster City, CA, USA) using the 310 Data Collection version 3.0.0. on a POP-6™ polymer. After the run, samples were analysed using GeneScan 3.7 software (Applied Biosystems).

DNA sequencing

To confirm the mutations identified by the SNaPShot kit, a random set of samples were sequenced with the above mentioned ABI 310 Genetic analyser (Applied Biosystems, Foster City, CA, USA), using ABI Prism Big-Dye Terminator Cycle Sequencing Ready Reaction Kit, version 1.1 (Applied Biosystems, Foster City, CA, USA). Sequences were aligned with the Chromas Lite program version 2.01 (http://www.technelysium.com.au/chromas_lite.html).

Samples

The SNaPshot system was validated by genotyping a set of samples from the Spanish population. As a part of a wider project, a total of 284 unrelated females of fertile age were selected. Blood samples were collected by venipuncture. Haematological parameters were determined following standard laboratory techniques and using the Symex NE 9100 automated haematology (Symex, Kobe, Japan) and the Modular Analytics Serum Work Area (Roche, Basel, Switzerland) analyzers. Measurements of haemoglobin concentration (Hb), serum ferritin (Ft), haematocrit (Hct), mean corpuscular volume (MCV), serum transferrin (Tf) and serum iron (Fe) were carried out and total iron binding capacity (TIBC (µmol/L) = 25.1 x Tf (g/L) and transferrin saturation x 100 and were calculated. The total sample was classified in three groups according to haemoglobin (g/dl) and ferritin

($\mu\text{g/l}$) values: control ($\text{Hb} \geq 12$, $\text{Ft} \geq 20$), deficient ($\text{Hb} \geq 12$ and $\text{Ft} < 20$ or $\text{Hb} \leq 12$ and $\text{Ft} > 20$) and anemic ($\text{Hb} < 12$ and $\text{Ft} < 20$).

DNA was isolated from peripheral blood by standard phenol chloroform protocol with proteinase K (adapted from Sambrook *et al.*, 1989). DNA samples have been quantified with the Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific USA).

The Ethics Committee of Hospital Clinica Puerta de Hierro and Consejo Superior de Investigaciones Cientificas, in Madrid (Spain) approved this study, and all participants gave their informed written consent.

Statistical analysis

Chi-squared test was used to validate multiplex results comparing observed genotype frequencies of each SNP in the 3 groups from the Spanish population with expected ones. $P < 0.05$ was considered statistically significant. Heterozygosity values and genotype frequencies of the 3 groups are reported in Table 2. The results and comparisons of genotype frequencies between group samples and between them and Hapmap data were performed according to Nei 1987 and by GenePop software (version 4.0.10) (Raymond and Rousset 1995; Rousset 2008).

TABLE 2: Genotypes and Heterozygosity Values

Group	SNPs	rs4820268	rs855791	rs1799852	rs2280673	rs1800562	rs3811647	rs2673289	rs1375515	rs1799945	rs16826756
1N: 30	AA	0.267±0.082			0.100±0.056	0.000±0.000	0.133±0.063		0.300±0.085		0.600±0.091
	GG	0.200±0.074				0.967±0.033	0.300±0.085		0.267±0.082	0.033±0.033	0.100±0.056
	CC		0.300±0.085	0.700±0.085	0.467±0.093			0.400±0.091		0.667±0.088	
	TT		0.167±0.069	0.000±0.000				0.167±0.069			
	AG	0.533±0.093				0.033±0.033	0.567±0.092		0.433±0.092		0.300±0.085
	CT		0.533±0.093	0.300±0.085							
H ± SD	AC				0.433±0.092						
	GC								0.300±0.085		
		0.506±0.015	0.499±0.021	0.259±0.065	0.440±0.045	0.034±0.032	0.495±0.024	0.481±0.031	0.508±0.013	0.304±0.064	Hm ± SD 0.391 ± 0.049
	AA	0.292±0.040									
	GG	0.185±0.034			0.115±0.028	0.000±0.000	0.108±0.027		0.438±0.044		0.654±0.042
	CC		0.338±0.042	0.777±0.037	0.400±0.043	0.969±0.015	0.485±0.044	0.346±0.042	0.085±0.025	0.046±0.018	0.023±0.013
2N: 130	TT		0.169±0.033	0.000±0.000				0.169±0.033		0.654±0.042	
	AG	0.523±0.044				0.031±0.015	0.408±0.043		0.477±0.044		0.323±0.041
	CT		0.492±0.044	0.223±0.037				0.485±0.044			
	AC				0.485±0.044						
	GC								0.300±0.040		
		0.496±0.007	0.487±0.011	0.200±0.030	0.461±0.017	0.030±0.015	0.431±0.022	0.486±0.011	0.439±0.021	0.316±0.030	Hm ± SD 0.365 ± 0.049
3N: 124	AA	0.274±0.040									
	GG	0.234±0.038			0.161±0.033	0.000±0.000	0.073±0.023		0.484±0.045		0.669±0.042
	CC		0.306±0.042	0.758±0.039	0.331±0.042	0.887±0.029	0.419±0.045		0.081±0.025	0.032±0.016	0.032±0.016
	TT		0.226±0.038	0.000±0.000				0.403±0.044		0.694±0.042	
	AG	0.492±0.045				0.113±0.029	0.508±0.045	0.137±0.031			
	CT		0.468±0.045	0.242±0.039				0.460±0.045	0.435±0.045		0.298±0.041
H ± SD	AC				0.508±0.045						
	GC								0.274±0.040		
		0.500±0.005	0.491±0.010	0.209±0.031	0.472±0.015	0.078±0.023	0.405±0.025	0.494±0.009	0.416±0.024	0.296±0.031	Hm ± SD 0.371 ± 0.044
	AA										
	GG										
	CC										

Group: 1 (anemic), 2 (deficient) and 3 (control)
(H: heterozygosity, Hm: average heterozygosity, SD: standard deviation, N: number of samples).

Results and Discussion

In this study, a simple and robust minisequencing assay was developed to simultaneously type 10 bi-allelic markers, which present high heterozygosity in European populations groups (The HapMap Project).

During the optimization, a range of quantities of template DNA (from 400ng/μl to 2ng/μl) was tested to determine better results. Smaller amounts of input DNA yielded reliable results, but the peak heights were smaller as expected.

To minimize the presence of unspecific peaks and for a more balanced output, we tested different primer mix concentrations in both reactions and different annealing temperatures and number of cycles for the multiplex PCR reaction. The main modification, which eliminated unspecific peaks, was the increase in cycle number at the multiplex PCR (Figure 1).

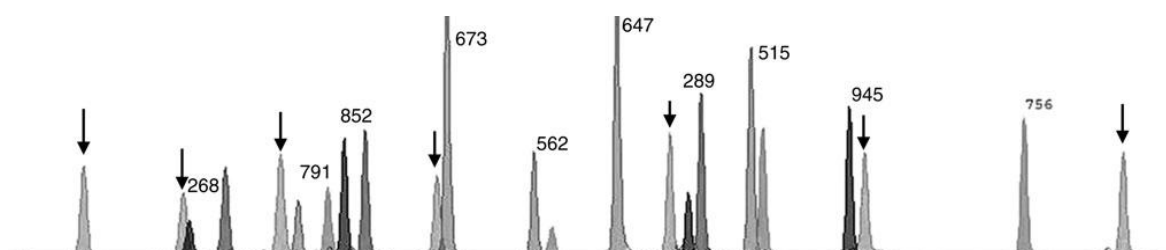


Figure 1. Results obtained from a SNaPshot. SNPs: 268 (rs4820268), 791 (rs855791), 852 (rs1799852), 673 (rs2280673), 562 (rs1800562), 647 (rs3811647), 289 (rs2673289), 515 (rs1375515), 945 (rs1799945), 756 (rs16826756). Red peaks represent T allele, Blue peaks represent G, green peaks represent A allele, Black peaks represent C allele, Orange peaks represent LIZ. Heterozygotes are 268, 791, 852, 562, 289, 515, and the rest of SNPs are homozygotes

In the multiplex PCR reaction that provides the templates for minisequencing, we increased the concentration of the weakly amplified products to equilibrate the resolution of the minisequencing output. Also for the SNaPshot reaction, it was necessary to experiment with different concentrations of each probe to obtain a balanced multiplex output. The poli-GACT tail at the 5' end proved to be more stable for longer single base extension primers (for example 71pb). As expected, the detected size determined by the automated sequencer and

the real size of certain products were slightly different. This is due to differences in electrophoretic mobility mainly determined by the length, sequence, and the dye used to label the extended primer. In some heterozygosity cases the effect was higher in shorter oligonucleotides than in longer ones.

The correct differentiation and identification of the extended primer was confirmed by sequencing the corresponding amplicon of the first PCR. SNaPshot and direct sequence results were in agreement in all cases. This Multiplex system confirmed the C282Y and H63D mutations in the HFE gene while these very mutations were previously detected in the same samples by restriction enzymes (Sarriá *et al.*, 2007a; Sarriá *et al.*, 2007b).

With respect to iron status among the 284 subjects, 124 were healthy (Hb ≥ 12 , Ft ≥ 20), 130 iron-deficient (Hb ≥ 12 and Ft < 20 or Hb ≤ 12 and Ft > 20) and 30 iron deficient anemic (Hb < 12 and Ft < 20).

The Chi-squared test confirms that all of the SNPs of the multiplex were in Hardy-Weinberg equilibrium in the 3 groups: group 1 ($\chi^2=5.8818$ P >0.9967), group 2 ($\chi^2=8.9376$ P >0.9836) and group 3 ($\chi^2=10.6582$ P >0.9546). Hardy-Weinberg equilibrium has been repeatedly recommended as a measure for quality control in the context of genetic association studies (Xu *et al.*, 2002). SNPs showing extreme departures of HWE in controls can be discarded, because this may indicate genotyping error, supposing that one of the assumptions of HWE does not fail (Wittke-Thompson *et al.*, 2005). Another quality of data parameter is to remove SNPs in which minor allele frequency (MAF) > 0.01 , in other words SNPs which are not polymorphic. In this study, each SNP presents MAF higher than 0.01.

The results of the screening in the 284 samples suggest a great degree of variability in the 3 groups: group 1 (Hm=0.391 \pm 0.049), group 2 (Hm=0.365 \pm 0.048) and group 3 (Hm=0.371 \pm 0.044) (Table 2). Moreover, all SNPs present high heterozygosity values except for rs1799852 and rs1800562.

As can be observed in table 2, there are differences in the genotype frequencies between group 1 (anemic) and group 3 (healthy), mainly in rs1375515, rs3811647 and rs2280673 SNPs. The higher frequency of rs2280673

AA genotype in healthy population decreases in the iron deficient group, reaching the lowest value in the anemic sample. On the contrary, the lowest frequency of rs3811647 AA genotype in healthy population increases in the anemic sample. This also happens in the case of rs1375515 with GG genotype. However, these differences were only statistically significant in rs1375515 between group 1 and group 2 ($P=0.024$), and group 3 ($P=0.009$). This is probably the most informative tag-SNP of the multiplex, being able to discriminate between the healthy and anemic groups. Furthermore, even if rs1800562 is the most homogeneous locus of the multiplex (Table 2), it is able to discriminate between the healthy and deficient groups ($P=0.014$).

Frequencies of genotypes of these 3 groups were compared locus by locus with genotypes obtained from population of European ancestry (CEPH collection) from International HapMap Project Database (The International HapMap Consortium *et al.*, 2003). SNP rs2673289 shows statistically significant differences between healthy women and data from HapMapNCBI ($p=0.037$). This result may be a population effect since the frequencies obtained in this study for this SNP could be characteristic of the Spanish population. In order to confirm this hypothesis, it will be necessary to analyze more data from other Iberian populations. Moreover, the sample from the CEPH collection is made up by men and women resident in Utah (USA) with Northern and Western European ancestry. In the present study, however, we only recollected women samples, so we do not rule out the possible effect of sex or population differences.

There are significant differences among the group of anemic and HapMap data for SNP rs1375515 ($P=0.012$). This is because the frequencies obtained in the GG genotype are 0.267 in the anemic group, 0.097 and 0.081 in the sample from HapMap and in the healthy group, respectively. This tag-SNP was selected from a particular region which showed a linkage signal with the levels of sTfR, detected by Remacha *et al.* (2006). It is necessary to conduct further population analysis to confirm its role as an indicator of changes in iron levels.

As described above, rs1800562 shows a significant differentiation between the deficient group and HapMap data ($P=0.036$). This result confirms the efficiency

of this SNPs in the genetic discrimination between the healthy and iron-deficient groups.

In conclusion, although the results obtained in this study indicate that it is necessary to find new more sensitive SNPs for the diagnosis of diseases related to iron status, the minisequencing-based methodology is suitable to be used as a tool for clinical diagnosis of iron metabolism disorders that are not easily classified as hemochromatosis or iron deficiency anemia. The method is very simple, sensitive and low cost. It is also a good option to replicate results obtained from GWAS (Remacha *et al.*, 2006; Benyamin *et al.*, 2009; Chambers *et al.*, 2009; Ganesh *et al.*, 2009; Soranzo *et al.*, 2009) and therefore represents an advance in the genetic diagnosis of SNP related with alterations of iron status.

Acknowledgement: We would like to thank Pedro Barrio and Gianmarco Ferri for technical support and advices. This study was supported by Spanish project AGL 2009-11437. S Bertoncini and R Blanco-Rojo were supported by Grupo Santander 2009 (Estancia doctores y tecnólogos UCM) and JAE-predoc grants, respectively.

References

- Bardien S, Human H, Harris T, et al. (2009) A rapid method for detection of five known mutations associated with aminoglycoside-induced deafness. *BMC Med Genet* 13;10:2.
- Benyamin B, McRae AF, Zhu G, et al. (2009) Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels. *Am J Hum Genet* 84:60-5.
- Bustamante-Aragones A, Gallego-Merlo J, Trujillo-Tiebas MJ, et al. (2008) New strategy for the prenatal detection / exclusion of paternal cystic fibrosis mutations in maternal plasma. *J Cyst Fibros* 7:505-10.
- Chambers JC, Zhang W, Li Y, et al. (2009) Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. *Nat Genet* 41:1170-2.
- Feder JN, Gnirke A, Thomas W, et al. (1996) A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* 13:399-408.

- Ferri G and Pelotti S (2009) Multiplex ABO genotyping by minisequencing. *Methods Mol Biol* 496:51-8.
- Filippini S, Blanco A, Fernández-Marmiesse A, et al. (2007) Multiplex SNaPshot for detection of BRCA1/2 common mutations in Spanish and Spanish related breast/ovarian cancer families. *BMC Med Genet* 29:8:40.
- Gabriel SB, Schaffner SF, Nguyen H, et al. (2002) The structure of haplotype blocks in the human genome. *Science* 21:2225-9.
- Ganesh SK, Zakai NA, van Rooij FJ, et al. (2009) Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nat Genet.* 41:1191-8.
- Navas-Carretero S, Baeza C, López-Parra AM et al. (2008) Ingesta de nutrientes y presencia de mutaciones en genes implicados en la deficiencia de hierro en un grupo de mujeres trabajadoras de la comunidad de Madrid. *Nutr Hosp* 1(2):144-145.
- Nei M (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.
- Patil N, Berno AJ, Hinds DA, et al. (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* 23:1719-23.
- Raymond M and Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Heredity* 86:248-249.
- Remacha AF, Souto JC, Soria JM, et al. (2006) Genomewide linkage analysis of soluble transferrin receptor plasma levels. *Ann Hematol.* 85:25-8.
- Rousset F (2008) Genepop'007: a complete reimplement of the Genepop software for Windows and Linux. *Mol Ecol Resources* 8: 103-106.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning. A laboratory manual*. In Sambrook J, Fritsch EF (eds) 2nd ed, New York: Cold Spring Harbour Laboratory Press.
- Sarria B, López-Parra AM, Navas-Carretero S, et al. (2007a) Hpcidin, Transferrin (exon 7) and Hemochromatosis genotyping suggests that haplotype block analysis is the best strategy for predicting iron deficiency phenotype in women. *Nutr Res* 27: 672-678.
- Sarriá B, Navas-Carretero S, Lopez-Parra AM, et al. (2007b) The G277S transferrin mutation does not affect iron absorption in iron deficient women. *Eur J Nutr* 46: 57-60.
- Soranzo N, Spector TD, Mangino M, et al. (2009) A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat Genet* 41:1182-90.
- Wittke-Thompson JK, Pluzhnikov A, Cox NJ (2005) Rational inferences about departures from Hardy-Weinberg equilibrium. *Am J Hum Genet* 76:967-86.

Stephens JC, Schneider JA, Tanguay DA, et al. (2001) Haplotype variation and linkage disequilibrium in 313 human genes. *Science* 20:489-93.

Xu J, Turner A, Little J, et al. (2002) Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Hum Genet* 111:573-4.

3. VARIABILIDAD DE LAS MUTACIONES RELACIONADAS CON EL METABOLISMO DEL HIERRO EN POBLACIÓN ESPAÑOLA

**3.1. “Variabilidad genética en estudios de asociación.
Estructura poblacional de 10 SNPs autosómicos
relacionados con el metabolismo del hierro en
mujeres españolas”**

2012

Publicado en:

BIODIVERSIDAD HUMANA Y EVOLUCIÓN, 1ª Edición - Noviembre 2012

Capítulo 5. Diversidad Genética de las Poblaciones Humanas: 413-6

Daniel Turbón Borrega, Lourdes Fañanás Saura,

Carme Rissech Badalló y Araceli Rosa de la Cruz

(editores)

Purpurink. Estudi editorial

ISBN (impreso): 978-84-695-6322-9

ISBN (PDF): 978-84-695-6323-6

Variabilidad genética en estudios de asociación. Estructura poblacional de 10 SNPs autosómicos relacionados con el metabolismo del hierro en mujeres españolas

Baeza Richer C¹, López-Parra AM¹, Blanco-Rojo R², Bertoncini S.³, Arroyo-Pardo E¹, Vaquero MP²

- 1 Departamento de Toxicología y Legislación Sanitaria, Facultad de Medicina, Universidad Complutense de Madrid. España.
- 2 Departamento de Metabolismo y Nutrición, ICTAN, (CSIC), Madrid, España.
- 3 Departamento de Biología, Unidad de Antropología, Universidad de Pisa, Pisa, Italia.

Autor de contacto: Carlos Baeza Richer.
Departamento de Toxicología y Legislación Sanitaria,
Facultad de Medicina.
Universidad Complutense de Madrid. España.
e-mail: cbaezaricher@med.ucm.es

RESUMEN

En general, en los estudios de asociación sólo se obtiene un número reducido de SNPs con p significativas ($p < 10^{-7}$). Técnicas como la minisequenciación pueden ser una excelente opción para confirmar o rechazar el interés de dichos SNPs y poder plantear su utilización en el campo del diagnóstico clínico y de la salud pública. Un total de 10 SNPs, fueron seleccionados según los resultados obtenidos en diferentes trabajos de genotipado masivo relacionados con el metabolismo del hierro. Se analizaron en 284 muestras de mujeres españolas no relacionadas, mediante una estrategia de *SNaPshot*. La muestra fue dividida en tres fenotipos relacionados con la absorción de hierro de acuerdo a sus valores de hemoglobina y ferritina: control, deficiente y anémico. Se calcularon distancias genéticas (F_{ST}) entre los tres grupos y con poblaciones europeas disponibles en el *NCBI Reference Assembly*. No se observaron distancias genéticas significativas entre todos los grupos comparados lo que confirma que los marcadores seleccionados no revelan ninguna estructuración poblacional. Por último, dos SNPs (rs1375515 y rs1800562) mostraron una distribución genotípica significativamente diferente entre los tres grupos fenotípicos, con lo que se discute su posible valor predictivo para la deficiencia del hierro.

PALABRAS CLAVE: SNPs, deficiencia de hierro, *SNaPshot*, estructura poblacional, estudios de asociación genética.

(GWAS) in relation with iron metabolism and were analysed using *SNaPshot* method in 284 not related Spanish women. The sample was divided into three groups according to their hemoglobin and ferritin values in order to classify different phenotypes in relation with their iron levels: control group, iron deficient and anaemic. Genetic distance (F_{ST}) was assessed among these groups and with European populations available in the NCBI Reference Assembly. No significant genetic distances were found confirming that the markers selected do not reveal any population substructure. Finally two SNPs (rs1375515, rs1800562) showed a genotypic distribution among the three phenotypic groups which was significantly different and its possible predictor value for iron deficiency is discussed.

KEY WORDS: SNPs, iron deficiency, *SNaPshot*, population structure, genetic association studies

INTRODUCCIÓN

Los estudios de asociación de SNPs han permitido un gran avance en la comprensión de la herencia compleja. Sin embargo, el análisis de un ingente de SNPs requiere un método discreto de selección de SNPs. Los SNPs seleccionados para estos estudios deben ser capaces de explicar la variación fenotípica y se exige replicar los resultados en donde es frecuente.

Biodiversidad humana y Evolución

Daniel Turbón Borrega
Lourdes Fañanás Saura
Carme Rissech Badalló
Araceli Rosa de la Cruz
(editores)



3.1 “Variabilidad genética en estudios de asociación. Estructura poblacional de 10 SNPs autosómicos relacionados con el metabolismo del hierro en mujeres españolas” Capítulo de libro

Baeza Richer C¹, López-Parra AM¹, Blanco-Rojo R², Bertoncini S.³, Arroyo-Pardo E¹, Vaquero MP²

¹Departamento de Toxicología y Legislación Sanitaria, Facultad de Medicina, Universidad Complutense de Madrid. España.

² Departamento de Metabolismo y Nutrición, ICTAN, (CSIC), Madrid, España.

³ Departamento de Biología, Unidad de Antropología, Universidad de Pisa, Pisa, Italia.

Autor de contacto: Carlos Baeza Richer, e-mail: cbaezaricher@med.ucm.es

Resumen

En general, en los estudios de asociación sólo se obtiene un número reducido de *SNPs* con *p* significativas ($p < 10^{-7}$). Técnicas como la minisecuencia pueden ser una excelente opción para confirmar o rechazar el interés de dichos *SNPs* y poder plantear su utilización en el campo del diagnóstico clínico y de la salud pública. Un total de 10 *SNPs*, fueron seleccionados según los resultados obtenidos en diferentes trabajos de genotipado masivo relacionados con el metabolismo del hierro. Se analizaron en 284 muestras de mujeres españolas no relacionadas, mediante una estrategia de *SNAPshot*. La muestra fue dividida en tres fenotipos relacionados con la absorción de hierro de acuerdo a sus valores de hemoglobina y ferritina: control, deficiente y anémico. Se calcularon distancias genéticas (*FST*) entre los tres grupos y con poblaciones europeas disponibles en el *NCBI Reference Assembly*. No se observaron distancias genéticas significativas entre todos los grupos comparados lo que confirma que los marcadores seleccionados no revelan ninguna estructuración poblacional. Por último, dos *SNPs* (rs1375515 y rs1800562) mostraron una distribución genotípica significativamente diferente entre los tres grupos fenotípicos, con lo que se discute su posible valor predictivo para la deficiencia del hierro.

Palabras clave: *SNPs, deficiencia de hierro, SNAPshot, estructura poblacional, estudios de asociación genética*

Title: Genetic variability in association studies. Population structure of 10 autosomal SNPs related with iron metabolism in Spanish women

Abstract

Generally, genetic association studies only obtain a small number of SNPs with high significative p-values ($p < 10^{-7}$). Thus, some analysis methods as the minisequencing may be an excellent option whether to confirm or reject the interest of this SNPs as well as to propose their usage both in the diagnostic and public health field. A set of ten SNPs were selected according to some Genome Association Studies (GWAS) in relation with iron metabolism and were analysed using SNaPhot method in 284 not related Spanish women. The sample was divided into three groups according to their hemoglobin and ferritin values in order to classify different phenotypes in relation with their iron levels: control group, iron deficient and anaemic. Genetic distance (F_{ST}) was assessed among these groups and with European populations available in the NCBI Reference Assembly. No significant genetic distances were found confirmig that the markers selected do not reveal any population substructure. Finally two SNPs (rs1375515, rs1800562) showed a genotypic distribution among the three phenotypic groups which was significantly different and its possible predictor value for iron deficiency is discussed.

Key words: *SNPs, iron deficiency, SNaPShot, population structure, genetic association studies*

Introducción

Los estudios de asociación mediante genotipado masivo han supuesto un gran avance técnico en la investigación de las enfermedades de herencia compleja. En dichos estudios aunque se maneja un número ingente de SNPs y otros marcadores, se suelen obtener un número discreto de *SNPs* asociados

significativamente ($p < 10^{-7}$) a determinados fenotipos o parámetros. Actualmente para dar validez científica a estos hallazgos y como criterio de calidad para su publicación se exige replicar el estudio en una muestra independiente y es aquí donde es frecuente encontrar dificultades (Chanock *et al*, 2007; Salas y Carracedo, 2007). Técnicas de desarrollo relativamente rápido y de bajo coste como la minisequenciación pueden ser una excelente opción para confirmar o rechazar (replicación) el interés de dichos *SNPs* y poder plantear su utilización en el campo del diagnóstico clínico y de la salud pública.

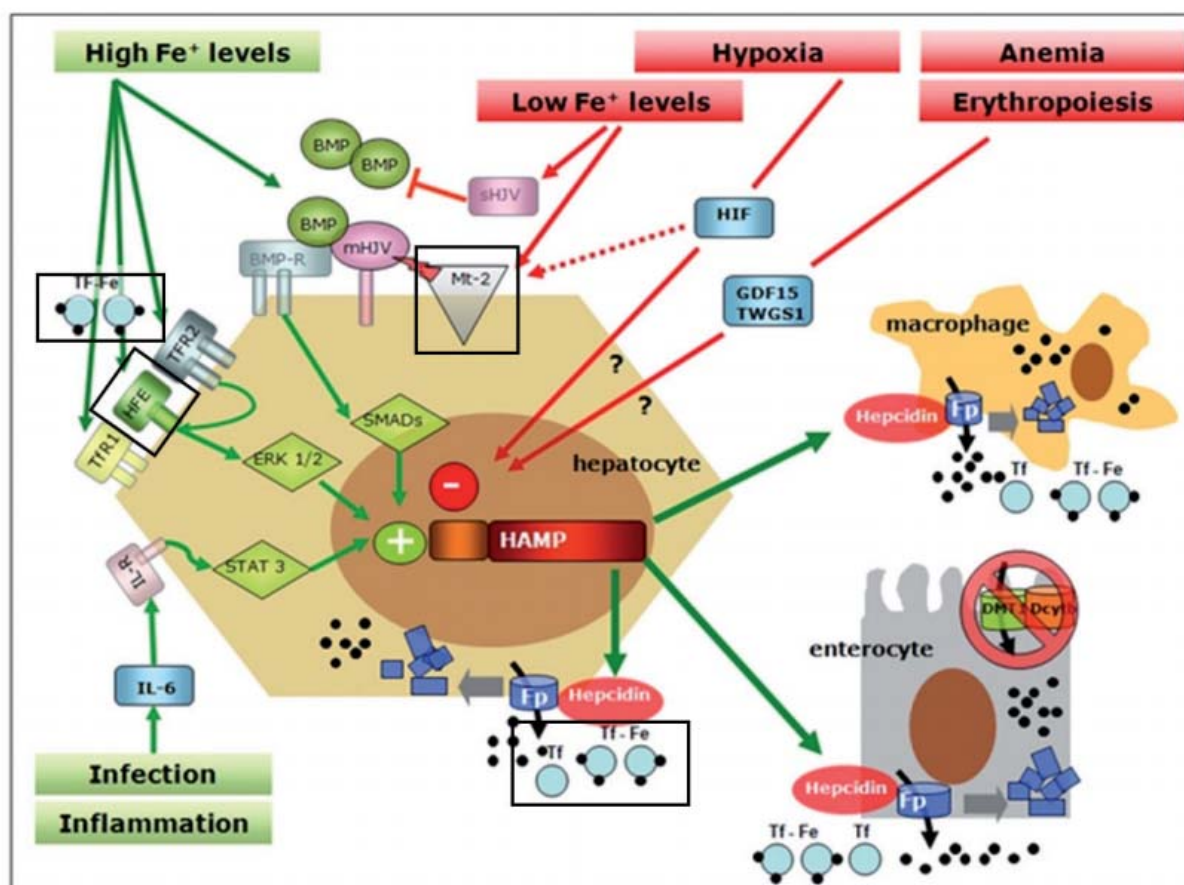


Figura 1. Esquema de la regulación del metabolismo del hierro. Modificado de Mariani *et al*, 2009

En el campo del metabolismo del hierro (Fig. 1) varios estudios de asociación mediante genotipado masivo (*GWAs*) han descubierto que ciertas mutaciones en proteínas clave de dicho metabolismo podrían ser factores de riesgo para enfermedades de herencia complejo como la hemocromatosis o la anemia siendo esta última una alteración calificada de pandemia por la

Organización Mundial de la Salud (OMS), que acarrea grandes costes económicos y sanitarios (de Benoist *et al*, 2008).

Una fuente de errores de tipo I y de tipo II en los GWAs es la selección de marcadores que presentan frecuencias muy diferenciadas según la población o bien la existencia de una fuerte subestructuración de las poblaciones analizadas (Balding, 2006; Salas y Carracedo, 2007). Así pues, los objetivos de este trabajo eran, por un lado, evaluar la técnica de la minisecuenciación como herramienta para replicar la asociación de 10 polimorfismos genéticos con parámetros relacionados con el metabolismo del hierro y, por otro lado, comparar nuestra muestra con otras poblaciones mundiales con el fin de revelar una posible estratificación poblacional para los mencionados polimorfismos.

Material y métodos

Marcadores seleccionados

Un total de 10 *SNPs* (rs4820268, rs855791, rs1799852, rs2280673, rs1800562, rs3811647, rs2673289, rs1375515, rs1799945, y rs16826756), fueron seleccionados según los resultados obtenidos en diferentes trabajos de genotipado masivo relacionados con el metabolismo del hierro (Feder *et al*, 1996; Remacha *et al*, 2006; Benyamin *et al*, 2009; Chambers *et al*, 2009). En la Figura 1 se pueden observar varias proteínas en recuadros negros (HFE, Matriptasa, Transferrina) implicadas en dicho metabolismo en cuyos genes se encuentran algunos de los *SNPs* estudiados. Estos marcadores se analizaron mediante la técnica *SNaPshot* cuyo protocolo diseñado específicamente para estos polimorfismos se describe en Bertoni *et al*, 2011. Un ejemplo de electroferograma obtenido mediante *SNaPshot* para los *SNPs* analizados se muestra en la Figura 2.

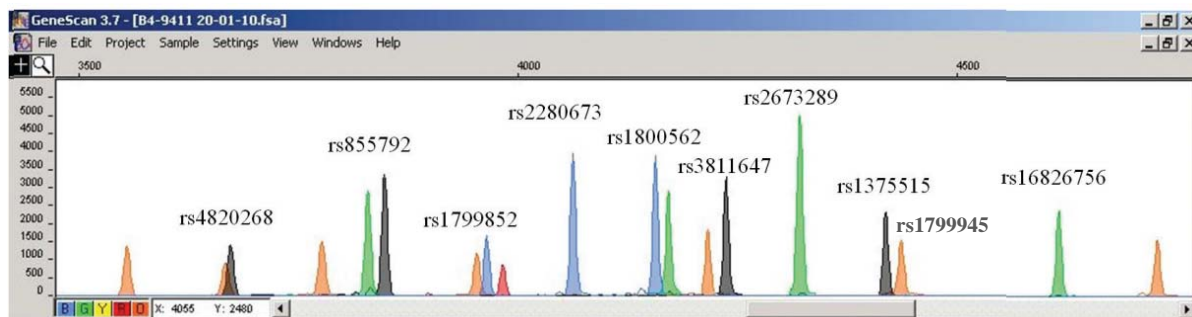


Figura 2. Electroferograma obtenido a partir de la técnica *SNaPshot* en la que se analizan 10 *SNPs* en una sola reacción

Muestra

Se analizaron en 284 muestras de mujeres españolas, no relacionadas. Esta muestra fue dividida en tres grupos de acuerdo a sus valores de hemoglobina (g/dL) y ferritina (mg/L) a fin de clasificar los diferentes fenotipos en relación a los niveles de hierro, según criterios de la OMS: control ($Hb > 12$, $Ft > 20$), deficiente ($Hb > 20$ y $Ft < 20$ o $Hb 12$ y $Ft > 20$), y anémico ($Hb < 12$ y $Ft < 20$).

Análisis estadísticos

Para testar la posible subestructuración, nuestra muestra, bien en conjunto o dividida en los tres grupos fenotípicos fue comparada con otros grupos poblacionales caucasoides empleando los datos disponibles en el *NCBI Reference Assembly* (www.ncbi.nlm.nih.gov/). Se calcularon las distancias genéticas entre los diferentes grupos en forma de *FST* mediante el paquete estadístico Arlequín ver. 3.5.1.2 (Excoffier *et al*, 2005). Por medio de un análisis chi-cuadrado, incluido en el paquete estadístico SPSS (versión 19), se analizaron las posibles diferencias en la distribución genotípica de los *SNPs* en los tres grupos fenotípicos.

Resultados y discusión

En el Figura 3 podemos observar la distribución de los diferentes genotipos de cada SNP en los 3 grupos fenotípicos. En algunos de los *SNPs* se aprecia una distribución desigual de genotipos entre los grupos. Así para el *SNP* rs1800562 se aprecia como la variante que produce la mutación C282Y, es más frecuente en normales que en el resto de los grupos. Este resultado es el esperado teniendo en cuenta que esta mutación en homocigosis presenta un riesgo muy elevado de generar hemocromatosis y por tanto sea asocia a mayores niveles de hierro (Feder *et al*, 1996). En otros *SNPs* también se aprecian distribuciones notablemente diferentes (rs1375515, rs3811647). De hecho, según los análisis de chi-cuadrado tanto para rs1375515 (p -valor: 0,026) como para rs1800562 (p -valor: 0,024) se encontraron diferencias significativas para la distribución genotípica en los tres grupos. En el caso de rs1375515, seleccionado a partir de una región que mostró una señal de ligamiento en el estudio de Remacha *et al*, 2006 se aprecia que el homocigoto mutante se distribuye preferentemente en el grupo de anémicas (Fig. 3). Estos resultados de chi-cuadrado sugieren que se debe profundizar en el estudio de rs1375515 y rs1800562 para comprobar su posible valor diagnóstico en la deficiencia del hierro.

El resto de *SNPs* no muestran diferencias significativas en sus distribuciones genotípicas. Esto se puede deber, entre otros motivos, a que no hay asociación entre estos *SNPs* y el desarrollo de la anemia o también a que los límites definidos para los grupos fenotípicos no sean los adecuados (Salas y Carracedo, 2007). En cualquier caso, la falta de significación no implica la ausencia de relación entre los *SNPs* y ciertos parámetros bioquímicos relacionados con el metabolismo del hierro tal y como se observa en otros estudios (Blanco-Rojo, en prensa).

A la hora de testar la posible subestructuración de la muestra, no se observaron distancias genéticas (F_{ST}) significativas para los *SNPs* analizados entre los 3 grupos y/o entre las diferentes poblaciones caucasoides disponibles (datos no mostrados). Por lo tanto no se observa estratificación poblacional para estos marcadores en población caucasoides.

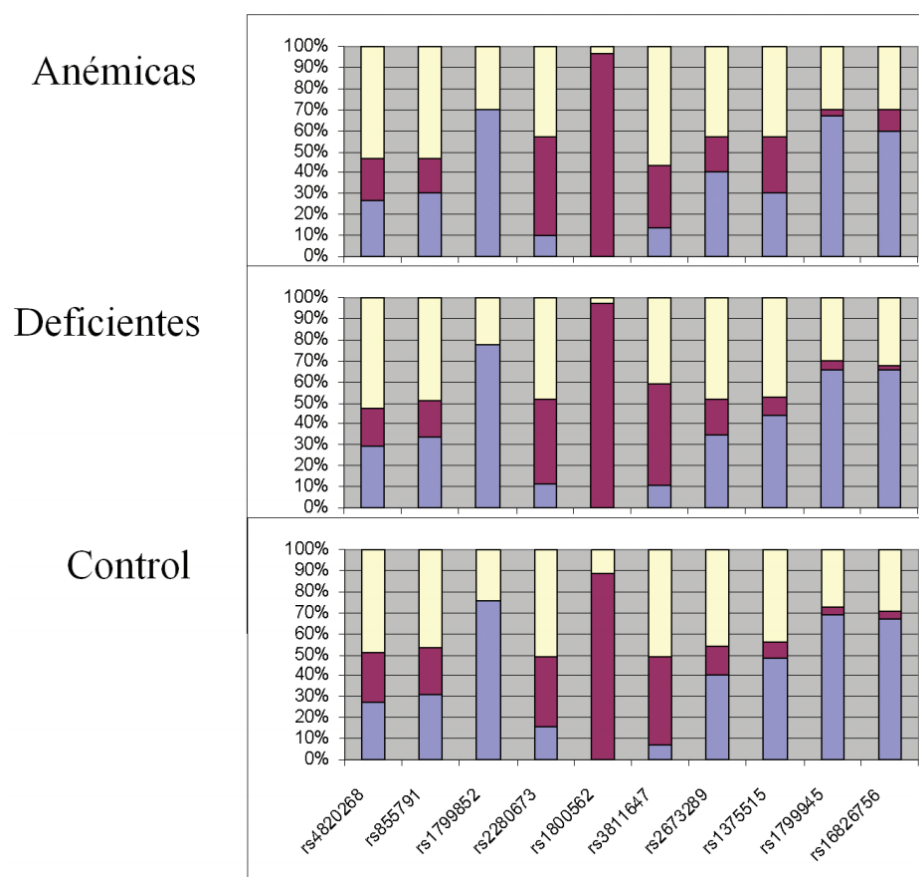


Figura 3. Distribución de los genotipos de cada SNP en los tres grupos fenotípicos. En el caso de rs1800562 el color amarillo representa el genotipo heterocigoto correspondiente a la mutación C282Y. En el caso de rs1375515 el color granate representa el homocigoto mutante perceptiblemente más abundante en el grupo de anémicas

En conclusión, a tenor de las diferentes distribuciones genotípicas encontradas entre los grupos y de la falta de estratificación poblacional, consideramos que este conjunto de *SNPs*, analizados mediante *SNaPshot*, puede ser de interés para replicar los resultados significativos encontrados en los estudios de asociación genética relacionados con metabolismo del hierro.

Bibliografía

Balding DJ (2006) A tutorial on statistical methods for population association studies. *Nat. Rev. Genet.*, 7(10):781-797.

- Benyamin B, Mcrae AF, Zhu G, Gordon S, Henders AK, Palotie A, Peltonen L, Martin NG, Montgomery GW, Whitfield JB, Visscher PM (2009). Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels. *Am. J. Hum. Genet.*, 84(1):60-65.
- Bertoncini S, Blanco-Rojo R, Baeza C, Arroyo-Pardo E, Vaquero MP Lopez-Parra AM (2011). A novel SNaPshot assay to detect genetic mutations related to iron metabolism. *Genet. Test. Mol. Biomarkers.*, 15(3):173-179.
- Blanco-Rojo R, Baeza-Richer C, López-Parra AM, Pérez-Granados AM, Brichs A, Bertoncini S, Buil A, Soria JM, Arroyo-Pardo E, Vaquero MP (2011) Four variants in transferrin and HFE genes as potential markers of iron deficiency anaemia risk: an association study in menstruating women. *Nutr. Metab.* (En prensa)
- Chambers JC, Zhang WLI Y, Sehmi J, Wass MN, Zabaneh D, Hoggart C, Bayele H, McCarthy MI, Peltonen L, Freimer NB, Srai SK, Maxwell PH, Sternberg MJ, Ruokonen A, Abecasis G, Jarvelin MR, Scott J, Elliott P, Kooner JS (2009) Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. *Nat. Genet.*, 41:1170-1172.
- Chanock SJ, Manolio T, Boehnke M (2007) Replicating genotype-phenotype associations. *Nature.*, 447 (7145):655-660.
- De Benoist B, MClean E, Egli I, Cogswell M (2008) Worldwide prevalence of anaemia 1993-2005: WHO Global Database on Anaemia. Geneva, World Health Organization..
- Excoffier L, Laval G, Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online.*, 1:47-50.
- Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A, Dormishian F, Domingo R Jr, Ellis MC, Fullan A, Hinton LM, Jones NL, Kimmel BE, Kronmal GS, Lauer P, Lee VK, Loeb DB, Mapa FA, McClelland E, Meyer NC, Mintier GA, Moeller N, Moore T, Morikang E, Prass CE, Quintana L, Starnes SM, Schatzman RC, Brunke KJ, Drayna DT, Risch NJ, Bacon BR, Wolff RK (1996). A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat. Genet.*, 13(4):399-408.
- Mariani R, Trombini P, Pozzi M, Piperno A (2009). Iron Metabolism in Thalassemia and Sick Cell Disease. *Mediterr J Hematol Infect Dis* 1: e2009006 DOI 10.4084/MJHID.2009.006.
- Remacha AF, Souto JC, Soria JM, Buil A, Sarda MP, Lathrop M, Blangero J, Almasy L, Fontcuberta J (2006). Genomewide linkage analysis of soluble transferrin receptor plasma levels. *Ann. Hematol.*, 85(1):25-28.
- Salas A, Carracedo A. (2007). Estudios de asociación en enfermedades complejas: problemas estadísticos relacionados con el análisis de polimorfismos genéticos. *Rev Clin Esp* 11:563-5.
- SPSS Inc. (1998). SPSS Base 8.0 for Windows User's Guide. SPSS Inc., Chicago. www.ncbi.nlm.nih.gov/

**3.2. “HFE, Hemojuvelin and Hepcidine Genotyping
Confirms High Prevalence of non-HFE Iron
Overload in Southern European Population”**

Bajo consideración editorial en:

Hum Biol

ISSN 0018-7143

3.2 “HFE, Hemojuvelin and Hepcidine Genotyping Confirms High Prevalence of non-HFE Iron Overload in Southern European Population” Artículo

C. Baeza-Richer¹, P. Ropero², F.A. González², M.S. Mesa³, A.M. López-Parra¹, A. Villegas², M.P. Vaquero⁴ and E. Arroyo-Pardo¹

¹ Department of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Madrid, Spain

² Service of Hematology and Hemotherapy, San Carlos Clinical Hospital. Madrid, Spain

³ Department of Zoology and Anthropology, Faculty of Biology, Complutense University of Madrid, Madrid, Spain

⁴ Department of Metabolism and Nutrition, Institute of Food Science, Food Technology and Nutrition, Spanish National Research Council, Madrid, Spain

Address for correspondence: Carlos Baeza Richer. Department of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Madrid, Spain.
cbaezaricher@med.ucm.es

Resumen

La sobrecarga de hierro está fuertemente relacionada con la hemocromatosis hereditaria, patología causada por mutaciones en proteínas clave del metabolismo del hierro. El tipo más común de esta enfermedad está causada por la mutación C282Y en homocigosis, localizada en el gen *HFE*, y cuya frecuencia es más elevada en poblaciones del norte de Europa. Mutaciones como la H63D en *HFE*, y otras localizadas en genes como la hemojuvelina (*HJV*) y la hepcidina (*HAMP*), parecen modular la severidad de esta enfermedad. Las hemocromatosis no causadas por el gen *HFE* son enfermedades raras debidas, en algunos casos, a alteraciones en *HJV*, *HAMP*, y otros genes. Este trabajo tiene como objetivo explorar la presencia de mutaciones en hepcidina y hemojuvelina en pacientes de población española con diferentes genotipos para *HFE* (C282Y y H63D). Para ello, se estudió una cohorte de 116 individuos de población española autóctona con sobrecarga de hierro. A su vez, se analizaron parámetros bioquímicos tales como ferritina, transferrina, la saturación de la transferrina, hemoglobina, volumen corpuscular medio y hematocrito. Los tests genéticos incluyeron el análisis de las regiones exónicas de

los genes *HAMP* y *HJV* mediante las técnicas SSCP y secuenciación, respectivamente, y el genotipado de las mutaciones del gen *HFE*: C282Y y H63D. Los resultados no mostraron ninguna mutación en *HJV* ni en *HAMP*, ni tampoco ninguna variación significativa en los parámetros bioquímicos en relación a los genotipos de *HFE*. Además, 30 individuos con sobrecarga no presentaron ninguna mutación para los tres genes estudiados. En conclusión, este estudio evidencia que las regiones exónicas de los genes *HJV* y *HAMP* raramente se encuentran involucrados en el desarrollo de la hemocromatosis en población española. Estos resultados confirman una prevalencia más baja de hemocromatosis debido a *HFE* en poblaciones mediterráneas.

Palabras clave: *sobrecarga de hierro primaria, hemocromatosis tipo no HFE, Hepsidina, Hemojuvelina, población mediterránea, frecuencia de C282Y*

Abstract

Iron overload is related with hereditary hemochromatosis, a genetic disorder caused by mutations in key proteins of the iron metabolism pathway. The most common is caused by the C282Y mutation when homozygous (*HFE* gene). C282Y is more present in Northern than in Southern Europe. H63D in *HFE* and mutations in other genes, such as hemojuvelin (*HJV*) and hepcidin (*HAMP*), can modulate the severity of the disease. Non-*HFE* hemochromatosis are rare diseases caused, in some cases, by alterations in *HJV*, *HAMP* or other genes. This study aims to explore the presence of mutations in hepcidin and hemojuvelin in patients with different genotypes of *HFE* gene (C282Y and H63D) in Spanish population. A cohort of 116 individuals with primary iron overload, with four grandparents belonging to Spanish population was analyzed. Biochemical parameters, ferritin, transferrin, haemoglobin, mean cell volume and pack cell volume were assessed. Genetic tests include the analysis of the exonic regions of *HAMP* and *HJV* and the *HFE* mutations (C282Y and H63D). The results show no mutation in *HJV* and *HAMP* and no variance in the biochemical measures among genotypes. In addition, 30 individuals lacked of any mutations in all the studied genes (*HFE*, *HJV* and *HAMP*). In

conclusion, this study evidences the rare involvement *HJV* and *HAMP* exonic regions in hemochromatosis in Spain. These findings also confirm a lower prevalence of *HFE* hemochromatosis in Mediterranean populations.

Key words: *primary iron overload, Non-HFE hemochromatosis, Hpcidin, Hemojuvelin, Mediterranean population, C282Y frequency*

Introduction

Iron overload is a physiological condition in which certain iron related parameters show elevated levels. Serum ferritin and transferrin saturation (TfS) are the most common parameters employed to describe this condition. Primary iron overload is defined when high iron parameters are not caused by other conditions, such as chronic liver diseases, viral hepatitis, alcoholic and nonalcoholic disease, hematologic processes with ineffective erythropoiesis, hemolytic anemias, and transfusional iron overload. In the Spanish population, the prevalence of general iron overload was found to be of 9.3% (Altès et al. 2004), consequently the prevalence of primary iron overload must be even rarer (Evangelista et al. 2015).

Primary iron overload is the most significant characteristic of hereditary hemochromatosis (HH). HH is a common genetic disorder of iron metabolism characterized by progressive accumulation of iron in parenchymal cells of multiple organs (Radio et al. 2014). The genetic basis of HH type I, the most common hemochromatosis, is well established and consists on different mutations in *HFE* (C282Y, and H63D to a lesser extent), a key protein that regulates iron metabolism (Feder et al. 1996, Pietrangelo 2004). The mutation C282Y when homozygous accounts for 80-90% of the HH cases in Northern Europe. HH type II is a rare disease linked with mutations in hemojuvelin (type IIa), and in hepcidin (type IIb) as described elsewhere (Papanikolaou et al. 2004; Roetto et al. 2004; Bayele et al. 2009). Regarding hemojuvelin (*HJV*) and hepcidin (*HAMP*), the exonic regions are the most interesting to be analyzed as virtually all mutations causing

haemochromatosis are located in these regions (Papanikolau et al. 2004; Brakensiek et al. 2009; Altès et al. 2009; Santos et al. 2010).

HFE protein and hemojuvelin are involved in iron metabolism by inducing hepcidin expression in high iron conditions (Dunn et al. 2007; Abbaspour et al. 2014). Hepcidin prevents iron from being absorbed in the enterocytes, stored in the hepatocytes and recycled from the macrophages. Therefore all the mutations that inhibit hepcidin expression or functionality are a potential risk of iron overload (Ganz and Nemeth 2012). Furthermore, the C282Y, when heterozygous, is reported to be protective of iron deficiency (Blanco-Rojo et al. 2014; Baeza-Richer et al. 2015). Our group already found no relationship between *HAMP* and iron deficiency (Sarria et al. 2007).

In the field of Anthropological Genetics, the study of geographic distribution of genetic variants has been employed to reconstruct human history and to identify adaptive and phenotypic variation (Eaaswarkhanth et al. 2014). In this line, C282Y mutations are more frequent in Northern than Southern Europe (Álvarez et al. 2001; Merryweather-Clarke et al. 2003). More specifically, in the Mediterranean region the proportion of hemochromatosis not explained by C282Y mutation is significantly higher, and other genes have been involved (Radio et al. 2014). Moreover, the penetrance in HH of C282Y is not complete (Rossi et al. 2008). Recently, some SNPs in the transferrin gene and others (*ARNTL*, *TFR2*) have been showed to be modifiers of iron metabolism in C282Y homozygous individuals (Benyamin et al. 2014; de Tayrac et al. 2015). Therefore, it is interesting to explore not only the *HFE* gene but also other genes related to iron metabolism, especially in Mediterranean populations.

The aim of this study is to assess the presence of mutations in hepcidin and hemojuvelin that could modulate the severity and/or etiology of primary iron overload in samples with different genotypes of *HFE* gene (C282Y and H63D) in Spanish population.

Material and methods

A cohort of 116 Spanish patients (70 men and 46 women) diagnosed with primary iron overload ($>400\mu\text{g/l}$ for ferritin and $>50\%$ for transferrin saturation), with or without symptoms of HH, and with no other reported disease was collected in the “Hospital Clinico San Carlos” of Madrid. DNA was extracted from fast blood samples using the method described by Miller et al. (1988).

Mutations (C282Y and H63D), from *HFE* gene were analysed by PCR using primers already designed by Feder et al. (1996) and Aguilar Martínez et al. (1997). The resulting two fragments were digested by *RsaI* (C282Y) and *DpnII* (H63D), and the products were visualized by electrophoresis on an 8% Acrylamide/Bis gel (Bio Rad laboratories, CA, USA), stained with ethidium bromide.

The exploration of mutations in the three exons of *HAMP* gene was carried by PCR amplifying exon 2 and 3 together as described by Zaahl et al. (2004). Amplicons were analyzed by single-strand conformational polymorphism (SSCP) in an ALF Sequencer (Pharmacia Biotech, Uppsala, Sweden), according to the manufacturer's protocol with slight modifications (ALFexpress, Pharmacia Biotech).

The four exons of *HJV* were analysed by sequencing. The first three exons were sequenced following Lanzara et al. (2004) strategy, introducing some changes in the PCR conditions. Aiming to sequence exon four, new primers were designed in order to obtain fragments smaller than 600 pb. Table 1 shows new primers sequences as well as the size of the resulting amplicons. PRIMER 3 software (Untergasser et al. 2012) was used to design four pair of primers to amplify four overlapping sequences (http://biotools.umassmed.edu/bioapps/primer3_www.cgi). PCR products were purified using QIAquick PCR purification kit (QIAGEN), and then sequenced using Big Dye Terminator Kit v1.1 cycle sequencing kit (Applied Biosystems) with the same PCR primers. Electrophoresis and detection were performed on an ABI PRISM 310 Genetic Analyzer. *HAMP* was studied for all the samples whereas *HJV* was studied in the group with no *HFE* mutations in order to find the possible

genetic basis of their overload. *HJV* was also studied in the group of samples which were heterozygous for both H63D and C282Y. This latter group is interesting because the penetrance of HH for this genotype is low and very variable among populations and between males and females, which suggests the existence of other genetic modifiers (Pereira et al. 2001; EASL et al. 2010; Gallego et al. 2015).

Patient haematological data [haemoglobin (Hb), packed cell volume (PCV) and mean corpuscular volume (MCV)] were obtained with an automated cell counter (Coulter LH750 Analyzer; Beckman Coulter, Brea, CA, USA). Biochemical parameters [ferritin, transferrin and transferrin saturation (TfS)] were studied by immunoassay systems (Access 2 Immunoassay System; Beckman Coulter, Brea, CA, USA).

Multifactorial ANOVA, using SPSS v22, was carried out to test the variation of the studied biochemical parameters regarding sex and *HFE* genotypes, followed by Bonferroni post hoc test.

Table 1. Oligonucleotides used for Polymerase Chain Reaction

Exon	Primer sequence (5' – 3')	Amplicon size	Reference
1F	CCGTCAACTCAGTAGCCA	336 bp	Lanzara et al. 2004
1R	CATTTGGACGAGAGACATC		
2F	CCCCAAATTCCAGTCTGTT	380 bp	Lanzara et al. 2004
2R	CTCATTCAGGCTCACATGC		
3F	CTCCGATAGAGCAGAGGTC	709 bp	Lanzara et al. 2004
3R	GGAGCATTGCTGTTGAATAG		
4a'F	TCAAGGATTGAGGGCCATAG	504 bp	Present study
4a'R	CGTCTGGCAGTATCAATGGTT		
4b'F	GCAGCTCTCCTTCTCCATCA	545 bp	Present study
4b'R	CAGCCTCATCTGACTCTGGA		
4c'F	TTCTGTGGCTTTGCATTGAG	408 bp	Present study
4c'R	TCATGTCTTCTGCTTTCAGCTC		
4d'F	TCCATTTACTGCAGATTTTACC	408 bp	Present study
4bR	CACATTGCATTTACTCCATTTC		Lanzara et al. 2004

Results and Discussion

The cohort of 116 Spanish patients suffering from primary iron overload is clustered according to *HFE* genotypes. This is a notable sample giving the low prevalence of primary iron overload in the Spanish population. Biochemical results for Hb, PCV, ferritin, MCV, transferrin and TfS in the four genetic clusters are shown in Table 2. There were no significant differences in the levels of any of the biochemical parameters regarding the defined groups by genotypes (Table 3). These results are noteworthy because when analysing general population, *HFE* genotypes have been significantly associated with high serum ferritin and transferrin saturation levels (McLaren and Gordeuk et al. 2009). In the present study, this association is not found likely because, as we have selected exclusively primary iron overload samples (a very specific subgroup of the population), other factors could also be involved in this condition. Besides, no mutation was found in hepcidin and hemojuvelin genes (*HAMP* and *HJV*) in any of the four clusters. This suggests that other mutations, possibly in other than *HJV* and *HAMP* are producing iron overload especially in the group with no C282Y and in the group with no mutations. In this line, we found 30 individuals (25.9% of the sample) in which iron overload could not be explained by mutations in any of the studied genes.

Present results support that hepcidin and hemojuvelin genes are conserved to a high extent, at least the exonic regions, and their variants are rare in Spanish populations as reported elsewhere in individuals homozygous for C282Y (Altès et al. 2009). This study shows that the exonic regions of *HJV* and *HAMP* do not explain neither the different penetrance of *HFE* iron overload nor the etiology of primary iron overload. The regulation of most of the iron metabolism pathways is mediated at some point by hepcidin (Abbaspour et al. 2014). Furthermore, hemojuvelin has been described as a key regulator of hepcidin (Zhang 2010), thus, their critical role would make this regions highly conserved due to natural selection. In fact, the majority of mutations are rare or private as shown in Table 4. Moreover, highly scarce information regarding population frequencies is available for these mutations. Mutation A197G in hemojuvelin is relatively frequent but its implication in iron overload is still discussed (Table 4). However, mutation in non-

Table 2. Genetic and biochemical characteristics of the cohort of 116 patients with iron overload

HFE	C282Y	H63D	HJV	HAMP	n	%	Hemoglobin (g/dL)			PCV (%)			Ferritin (ng/mL)			MCV (%)			Transferrin (mg/dL)			SAT (%)		
							Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
-/-	-/-	Wild Type	Wild Type	Wild Type	30	25.9	14.877	0.353	44.300	1.071	875.400	125.993	95.517	0.961	244.667	12.546	55.333	26.975						
-/-	+/+	Wild Type	Wild Type	Wild Type	31	26.7	14.097	0.225	40.752	0.673	473.632	106.620	93.858	1.091	248.355	16.067	50.548	28.266						
-/+	-/+	Wild Type	Wild Type	Wild Type	31	26.7	14.923	0.294	43.813	0.762	455.071	60.125	94.119	0.755	267.774	9.0205	67.710	27.815						
+/+	-/-	Wild Type	Wild Type	Wild Type	24	20.7	14.342	0.319	40.783	1.794	554.125	136.448	92.962	1.340	224.500	12.684	53.375	23.279						
Total					116	100.0	14.570	1.634	42.494	5.936	589.231	600.584	94.172	5.536	247.655	70.425	56.957	27.345						

-/- (wild type homozygous), +/- (heterozygous), +/+ (mutant homozygous)

Table 3. Multifactorial ANOVAs for each biochemical parameter, as dependent variable, and genotype and sex, as factors

ANOVA	Hemoglobin			PCV			Ferritin			MCV			Transferrin			SAT		
	F	p	F	F	p	F	F	p	F	F	p	F	F	p	F	F	p	F
Genotype	0.621	0.603	1.788	0.154	2.319	0.079	1.041	0.377	1.812	0.149	2.503	0.063						
Sex	11.595	0.001	2.311	0.131	3.504	0.064	0.176	0.676	0.195	0.66	0.281	0.597						

coding regions should not be discarded (Bayele et al. 2006 and 2009; Island et al. 2009).

Table 4. Worldwide main mutations of *HFE*, *HJV* and *HAMP* genes
– Frequencies at population level not available

Gene	Mutations	Allele Frequency	Population or ethnicity	Reference
HFE	C282Y	G(0.947) A(0.053)	Northern, Western European	CEU, HapMap
		G(0.972) A(0.028)	Italian	TSI, HapMap
		G(0.980) A(0.020)	Spanish	Álvarez et al. 2001
HJV (HFE2)	G320V	G(0.999) T(0.001)	North American	NCBI database
		—	French Canadian	Lanzara et al. 2004
		—	Greek Families	Papanikolau et al. 2004
		—	Italian/Canadian	Lanzara et al. 2004
		—	European	Lee et al. 2004a
		—	Central European	Gehrke et al. 2005
		—	German	Brakensiek et al. 2009
	A197G	C(0.898) G(0.102)	Yoruban, Nigeria	YRI, HapMap
		C(0.896) G(0.104)	African Ancestry	ASW, HapMap
		C(0.990) G(0.010)	Mexican Ancestry	MEX, HapMap
	S264L	—	Brazilian	Santos et al. 2010
		—	Spanish Family	Altès et al. 2009
		—	North American	Barton et al. 2002
	C80R	—	North American	Lee et al. 2004b
		—	North American	Lee et al. 2004a
	C321X	—	Chinese	Huang et al. 2004
		—	Greek Families	Papanikolau et al. 2004
	I281T	—	Chinese	Huang et al. 2004
		—	African American	Murugan et al. 2008
HAMP	G71D	G(0.997) A(0.003)	North American	NCBI database
		—	European family	Merryweather-Clarke et al. 2003
		—	Spanish Family	Altès et al. 2009
	R56X	—	Hispanic	Barton et al. 2009
		—	Italian family	Roetto et al. 2003
		—	Greek Family	Jacolot et al. 2004
	Y70R	—	Italian	Roetto et al. 2004
		—	Italian	Majore et al. 2004
	C78Y	—	Australian	Delatycki et al. 2004
	R59G	—	North American	Jacolot et al. 2004
		—	French	Barton et al. 2009

The C282Y mutation is believed to be of Celtic origin (Merryweather-Clarke et al. 2003) and has a higher distribution in Northern Europe presenting a North West, South East descendant cline as shown in Table 4 (Campbell et al. 2003;

Merryweather-Clarke et al. 2003; Milić et al. 2011). Specifically in Spain this pattern holds as the frequencies decrease from north to south (Álvarez et al. 2001; Ropero-Gradilla et al. 2005). Therefore, as current study confirms, higher proportion of non-*HFE* iron overload is expected in Southern populations. In that respect, Radio et al. (2014) found a significant proportion of non-*HFE* hemochromatosis in an Italian cohort. In that case several *TfR2* (transferrin receptor 2) mutations were found.

In conclusion, this study evidences that *HJV* and *HAMP* exonic regions are rarely involved in non-*HFE* hemochromatosis in primary iron overload Spanish individuals, and supports that the majority of mutations in those genes are rare and private. Finally, our findings confirm a lower prevalence of iron overload and hemochromatosis due to *HFE* in Mediterranean populations. This makes it necessary to explore in depth, employing larger sample sizes, the origin of this metabolic disorder in populations with low frequency of those mutations, such as the Spanish.

Acknowledgements

This study has been financed by FMM-05 research project.

References

- Abbaspour, N., Hurrell, R. and Kelishadi, R. 2014. Review on iron and its importance for human health. *J. Res. Med. Sci.* 19:164-74.
- Aguilar Martinez, P., Jeanjean, P., Masmejean, C., et al. 1997. Simple and rapid detection of the newly described mutations in the HLA-H gene. *Blood*. 89:1835-6.
- ALF Express DNA Sequencer*. 1992. User Manual, Pharmacia Biotech. Uppsala, Sweden.
- Altès, A., Bach, V., Ruiz, A., et al. 2009. Mutations in *HAMP* and *HJV* genes and their impact on expression of clinical hemochromatosis in a cohort of 100 Spanish patients homozygous for the C282Y mutation of *HFE* gene. *Ann. Hematol.* 88:951-5.
- Altès, A., Ruiza, MA., Castell, C., et al. 2004. Déficit y sobrecarga de hierro en la población adulta de Cataluña. *Med. Clin. (Barc)*. 123(4):131-3.
- Álvarez, S., Mesa, MS., Bandrés, F., et al. 2001. C282Y and H63D mutation frequencies in a population from Central Spain. *Dis. Markers*. 17:111-4.

- Baeza-Richer C., Arroyo-Pardo, E., Blanco-Rojo, R., et al. 2015. Genetic contribution to iron status: SNPs related to iron deficiency anaemia and fine mapping of CACNA2D3 calcium channel subunit. *Blood Cells Mol. Dis.* 55:273-80.
- Barton, JC., Lafreniere, SA., Leiendecker-Foster, C., et al. 2009. HFE, SLC40A1, HAMP, HJV, TFR2, and FTL mutations detected by denaturing high-performance liquid chromatography after iron phenotyping and HFE C282Y and H63D genotyping in 785 HEIRS Study participants. *Am. J. Hematol.* 84:710-4.
- Barton, JC., Rao, SV., Pereira, NM., et al. 2002. Juvenile hemochromatosis in the Southeastern United States: a report of seven cases in two kinships. *Blood. Cells. Molec. Dis.* 29:104-5.
- Bayele, HK. and Srai, SK. 2009. Genetic Variation in Hepcidin Expression and its Implications for Phenotypic Differences In Iron Metabolism. *Haematologica.* 94:1185-8.
- Bayele, HK., McArdle, H. and Srai, SK. 2006. Cis and trans regulation of hepcidin expression by upstream stimulatory factor. *Blood.* 108:4237-45.
- Benyamin, B., Esko, T., Ried, JS., et al. 2014. Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis. *Nat Commun.* 5:4926.
- Blanco-Rojo, R., Toxqui, L., López-Parra, AM., et al. 2014. Influence of diet, menstruation and genetic factors on iron status: a cross-sectional study in Spanish women of childbearing age. *Int. J. Mol. Sci.* 15:4077-87.
- Brakensiek, K., Fegbeutel, C., Mälzer, M., et al. 2009. Juvenile hemochromatosis due to homozygosity for the G320V mutation in the HJV gene with fatal outcome. *Clin. Genet.* 76:493-5.
- Campbell, S., George, DK., Robb, SD., et al. 2003. The prevalence of haemochromatosis gene mutations in the West of Scotland and their relation to ischaemic heart disease. *Heart.* 89:1023-6.
- de Tayrac, M., Roth, MP., Jouanolle, AM., et al. 2015. Genome-wide association study identifies TF as a significant modifier gene of iron metabolism in HFE hemochromatosis. *J. Hepatol.* 62(3):664-72.
- Delatycki, MB., Allen, KJ., Gow, P., et al. 2004. A homozygous HAMP mutation in a multiply consanguineous family with pseudo-dominant juvenile hemochromatosis. *Clin. Genet.* 65:378-83.
- Dunn, LL., Suryo Rahmanto, Y. and Richardson, DR. 2007. Iron uptake and metabolism in the new millennium. *Trends Cell. Biol.* 17:93-100.
- Eaaswarkhanth, M., Pavlidis, P., and Gokcumen, O. Geographic distribution and adaptive significance of genomic structural variants: an anthropological genetics perspective. *Hum. Biol.* 2014. 86:260-75.
- Edwards, CQ., Ajioka, RS. and Kushner, JP. 2000. Hemochromatosis: a genetic definition. In *Hemochromatosis: Genetics, Pathophysiology, Diagnosis and Treatment*. Barton JC., Edwards CQ., eds. Cambridge, United Kingdom: Cambridge University Press, 8-11.
- European Association For The Study Of The Liver. 2010. EASL clinical practice guidelines for HFE hemochromatosis. *J. Hepatol.* 53(1):3-22.

- Evangelista, AS., Nakhle, MC., de Araújo, TF., et al. 2015. HFE genotyping in patients with elevated serum iron indices and liver diseases. *Biomed. Res. Int.*:164671.
- Feder, JN., Gnirke, A., Thomas, W., et al. 1996. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat. Genet.* 13:399-408
- Gallego, CJ., Burt, A., Sundaresan, AS., et al. 2015. Penetrance of Hemochromatosis in HFE Genotypes Resulting in p.Cys282Tyr and p.[Cys282Tyr];[His63Asp] in the eMERGE Network. *Am. J. Hum. Genet.* pii: S0002-9297(15)00330-4.
- Ganz, T. and Nemeth, E. 2012. Hepcidin and iron homeostasis. *Biochim. Biophys. Acta.* 1823:1434-43.
- Gehrke, SG., Pietrangelo, A., Kascák, M., et al. 2005. HJV gene mutations in European patients with juvenile hemochromatosis. *Clin. Genet.* 67:425-8.
- HapMap. The International HapMap Project. Available at: <http://hapmap.ncbi.nlm.nih.gov/> [Accessed sep 2015].
- Huang, FW., Rubio-Aliaga, I., Kushner, JP., et al. 2004. Identification of a novel mutation (C321X) in HJV. *Blood.* 104:2176-77.
- Island, ML., Jouanolle, AM., Mosser, A., et al. 2009. A new mutation in the hepcidin promoter impairs its BMP response and contributes to a severe phenotype in HFE related hemochromatosis. *Haematologica.* 94:720-4.
- Jacot, S., Le Gac, G., Scotet, V., et al. 2004. HAMP as a modifier gene that increases the phenotypic expression of the HFE pC282Y homozygous genotype. *Blood.* 103:2835-40.
- Lanzara, C., Roetto, A., Daraio, F., et al. 2004. Spectrum of hemojuvelin gene mutations in 1q-linked juvenile hemochromatosis. *Blood.* 103:4317-21.
- Lee, PL., Barton, JC., Brandhagen, D., et al. 2004a. Hemojuvelin (HJV) mutations in persons of European, African-American and Asian ancestry with adult onset haemochromatosis. *Br. J. Haematol.* 127:224-9.
- Lee, PL., Beutler, E., Rao, SV., et al. 2004b. Genetic abnormalities and juvenile hemochromatosis: mutations of the HJV gene encoding hemojuvelin. *Blood.* 103:4669-71.
- Majore, S., Binni, F., Pennese, A., et al. 2004. HAMP gene mutation c.208T>C (p.C70R) identified in an Italian patient with severe hereditary hemochromatosis. *Hum. Mutat.* 23:400.
- McLaren, GD. and Gordeuk, VR. 2009. Hereditary hemochromatosis: insights from the Hemochromatosis and Iron Overload Screening (HEIRS) Study. *Hematology Am. Soc. Hematol. Educ. Program*:195-206.
- Merryweather-Clarke, AT., Cadet, E., Bomford, A., et al. 2003. Digenic inheritance of mutations in HAMP and HFE results in different types of haemochromatosis. *Hum Mol Genet.* 12(17):2241-7.
- Milić, S., Ristić, S., Starčević-Čizmarević, N., et al. 2011. Low frequency of HFE gene mutations in Croatian patients suspected of having hereditary hemochromatosis. *Med. Sci. Monit.* 17(10):CR552-6.

- Miller, SA., Dykes, DD. and Polesky, HF. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic. Acids. Res.* 16:1215.
- Murugan, RC., Lee, PL., Kalavar, MR., et al. 2008. Early age-of-onset iron overload and homozygosity for the novel hemojuvelin mutation HJV R54X (exon 3; c.160A-->T) in an African American male of West Indies descent. *Clin. Genet.* 74:88-92.
- NCBI database. dbSNP Short Genetic Variations. Available at: <http://www.ncbi.nlm.nih.gov/projects/SNP/> [Accessed sep 2015].
- Papanikolaou, G., Sauels, ME., Ludwig, EH., et al. 2004. Mutation in HFE2 cause iron overload in chromosome 1q-linked juvenile hemochromatosis. *Nat Genet* 36:77–82.
- Pereira, AC., Mota, GF., and Krieger, JE. Hemochromatosis gene variants in three different ethnic populations: effects of admixture for screening programs. *Hum. Biol.* 2001. 73:145-51.
- Pietrangelo, A. 2004. Hereditary hemochromatosis - a new look at an old disease. *N. Engl. J. Med.* 350:2383-97.
- Radio, FC., Majore, S., Binni, F., et al. 2014. TFR2-related hereditary hemochromatosis as a frequent cause of primary iron overload in patients from Central-Southern Italy. *Blood. Cells. Mol. Dis.* 52:83-7.
- Roetto, A., Daraio, F., Porporato, P., et al. 2004. Screening hepcidin for mutations in juvenile hemochromatosis: identification of a new mutation (C70R). *Blood.* 103:2407–9.
- Roetto, A., Papanikolaou, G., Politou, M., et al. 2003. Mutant antimicrobial peptide hepcidin is associated with severe juvenile hemochromatosis. *Nat. Genet.* 33:21-2.
- Ropero-Gradilla, P., González-Fernández, FA., Briceño-Polacre, O., et al. 2005. Geographical distribution of HFE C282Y and H63D mutation in Spain. *Med. Clin. (Barc).* 125:794-5.
- Santos, PC., Cançado, RD., Pereira, AC., et al. 2010. HJV hemochromatosis, iron overload, and hypogonadism in a Brazilian man: treatment with phlebotomy and deferasirox. *Acta. Haematol.* 124:204-5.
- Sarriá, B., López-Parra, AM., Navas-Carretero, S., et al 2007. Hepcidin, transferrin (exon 7), and hemochromatosis genotyping suggests that haplotype block analysis is the best strategy for predicting iron deficiency phenotype in women. *Nutr. Res.* 27:672-8.
- Untergasser, A., Cutcutache, I., Koressaar, T., et al. 2012. Primer3 - new capabilities and interfaces. *Nucleic. Acids. Res.* 40:e115.
- Zaahl, MG., Merryweather-Clarke, AT., Kotze, MJ., et al. 2004. Analysis of genes implicated in iron regulation in individuals presenting with primary iron overload. *Hum. Genet.* 115:409-17.
- Zhang, AS. 2010. Control of Systemic Iron Homeostasis by the Hemojuvelin - Hepcidin Axis. *Adv. Nutr.* 1:38-45.

4. ESTUDIOS DE ASOCIACIÓN DE MARCADORES GENÉTICOS CON PARÁMETROS BIOQUÍMICOS Y FENOTIPOS CLÍNICOS RELACIONADOS CON EL METABOLISMO DEL HIERRO

**4.1. “Identification of a novel Quantitative Trait
Nucleotype related to iron status in a calcium
channel gene”**

2013

Dis Markers. 2013;34(2):121-9.

doi: 10.3233/DMA-120951

PMID: 23324578

Identification of a novel quantitative trait nucleotide related to iron status in a calcium channel gene

Carlos Baeza-Richer^{a,*}, Ruth Blanco-Rojó^b, Ana M. López-Parra^a, Anna Brichs^c, Stefania Bertoncini^d, Ana M. Pérez-Granados^b, Alfonso Buil^c, José M. Soria^c, Eduardo Arroyo-Pardo^a and M. Pilar Vaquero^b

^aDepartment of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Madrid, Spain

^bDepartment of Metabolism and Nutrition, Institute of Food Science, Food Technology and Nutrition, Spanish National Research Council, Madrid, Spain

^cUnit of Genomic of Complex Diseases, Institute of Biomedical Research, Barcelona, Spain

^dDepartment of Biology, University of Pisa, Pisa, Italy

Abstract. Several iron-related parameters have been reported to show significant heritability, and thus, seemed to be genetically regulated. A genome wide family-based study revealed two regions that showed a linkage signal with transferrin receptor levels. The aim of the study was to identify genetic markers associated with iron status biomarkers. Ten SNPs selected from the literature were tested, and parameters related to iron metabolism were analysed, in a group ($n = 284$) of Spanish women. Data were analyzed using Bayesian Model Averaging (BMA) test and decision trees. The rs1375515, located in an intronic region of the calcium channel gene *CACNA2D3*, showed strong associations with levels of mean corpuscular volume according to BMA test, and with levels of haemoglobin and ferritin according to decision trees. The allele G was associated to low levels of these parameters which suggests higher iron deficiency anaemia risk. This SNP along with the C282Y mutation explained significant differences in the distribution of individuals in three iron-related clinical phenotypes (normal, iron deficient and iron deficiency anaemic). In conclusion, the rs1375515, or other genetic polymorphisms in linkage, may play important roles in iron status, probably by affecting the function of a calcium channel. These findings may be useful for further investigation in the etiology of iron diseases.

Keywords: Quantitative trait nucleotide, iron deficiency anaemia, calcium channel gene, SNP, association study

1. Introduction

Specific levels of iron in biological fluids are essential for normal body function, in oxygen transport and for other important metabolic reactions. Thus, fine regulation of this element is required since departures from its optimal levels produce severe alterations. For example, iron overload triggers the Fenton reaction in

which the generation of hydroxyl radicals, causes may- or tissue damages [1]. By contrast, iron deficiency may cause anaemia. As in mammals there are no active mechanisms to eliminate iron, its absorption and storage regulation are crucial for maintaining appropriate levels of this mineral.

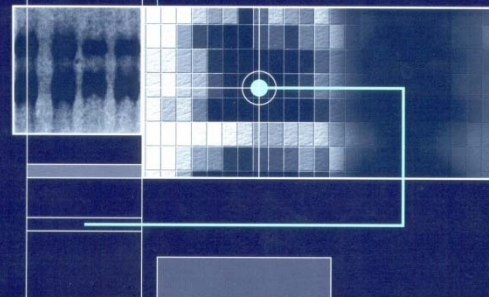
Among iron related disorders, iron deficiency is a worldwide world spread. According to WHO, one quarter of the world population is iron deficient. The prevalence of iron deficiency is higher in women than in men, being women at risk of iron deficiency.

*Corresponding author: Carlos Baeza-Richer, Department of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Complutense Avenue, 28040 Madrid, Spain. Tel: +34 91 394 11 57; E-mail: cbaeza@cclema.com

DISEASE MARKERS

Volume 17, Number 2

- ☐ DNA polymorphisms and mutations
- ☐ Biomarker proteins
- ☐ Clinical findings
- ☐ Cytogenetics
- ☐ Tissue antigens
- ☐ Histopathology
- ☐ Microbiology
- ☐ Epidemiology



IOS
Press
SPECIAL ISSUE: Cancer Genomics
Guest Editor: Robert L. Strausberg

ISSN 0278-0240

4.1 “Identification of a novel Quantitative Trait Nucleotype related to iron status in a calcium channel gene” Artículo

Carlos Baeza-Richer, ^{1*} Ruth Blanco-Rojo, ² Ana M. López-Parra, ¹ Anna Brichs, ³ Stefania Bertoncini, ⁴ Ana M. Pérez-Granados, ² Alfonso Buil, ³ José M. Soria, ³ Eduardo Arroyo-Pardo, ¹ and M. Pilar Vaquero ²

¹ Department of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Madrid, Spain

² Department of Metabolism and Nutrition, Institute of Food Science, Food Technology and Nutrition, Spanish National Research Council, Madrid, Spain

³ Unit of Genomic of Complex Diseases, Institute of Biomedical Research, Barcelona, Spain

⁴ Department of Biology, University of Pisa, Pisa, Italy

* Author for correspondence: Carlos Baeza Richer, e-mail: cbaezaricher@med.ucm.es

Resumen

Numerosos estudios demuestran como ciertos parámetros relacionados con el hierro presentan una heredabilidad alta y por tanto parecen estar regulados genéticamente. Además un estudio de ligamiento familiar reveló dos regiones ligadas a la variación de los niveles de receptor de la transferrina. El objetivo de este trabajo era identificar marcadores genéticos asociados a parámetros bioquímicos que se emplean para valorar el estado del hierro. Con este fin se analizaron 10 SNPs extraídos de la literatura, y se valoraron los parámetros relacionados con el hierro en un grupo (n= 284) de mujeres españolas. Los datos se analizaron mediante el test *Bayesian Model Averaging* (BMA), y mediante árboles de decisión. El rs1375515, localizado en una región intrónica del gen del canal del calcio *CACNA2D3*, se asoció fuertemente con los niveles de volumen corpuscular medio de acuerdo al BMA, y con los niveles de hemoglobina y ferritina según los árboles de decisión. El alelo G se asoció con niveles bajos de estos parámetros lo que sugiere un riesgo aumentado de sufrir anemia. Este SNP junto con la mutación C282Y explicaban diferencias significativas en la distribución de los individuos entre los fenotipos clínicos del hierro. En conclusión, el rs1375515,

u otros polimorfismos en ligamiento, pueden jugar un papel importante en la regulación del estado del hierro posiblemente al afectar la función del canal del calcio. Estos hallazgos podrían ser de utilidad en la investigación de la etiología de las patologías del hierro.

Palabras clave: *nucleótido de caracteres cuantitativos, anemia ferropénica, gen de canal del calcio, SNP, estudio de asociación.*

Abstract

Several iron-related parameters have been reported to show significant heritability, and thus, seemed to be genetically regulated. A genome wide family-based study revealed two regions that showed a linkage signal with transferrin receptor levels. The aim of the study was to identify genetic markers associated with iron status biomarkers. Ten SNPs selected from the literature were tested, and parameters related to iron metabolism were analysed, in a group (n=284) of Spanish women. Data were analyzed using Bayesian Model Averaging (BMA) test and decision trees. The rs1375515, located in an intronic region of the calcium channel gene *CACNA2D3*, showed strong associations with levels of mean corpuscular volume according to BMA test, and with levels of haemoglobin and ferritin according to decision trees. The allele G was associated to low levels of these parameters which suggests higher iron deficiency anaemia risk. This SNP along with the C282Y mutation explained significant differences in the distribution of individuals in three iron-related clinical phenotypes (normal, iron deficient and iron deficiency anaemic). In conclusion, the rs1375515, or other genetic polymorphisms in linkage, may play important roles in iron status, probably by affecting the function of a calcium channel. These findings may be useful for further investigation in the etiology of iron diseases.

Keywords: *quantitative trait nucleotide, iron deficiency anaemia, calcium channel gene, SNP, association study.*

1. Introduction

Specific levels of iron in biological fluids are essential for normal body function, in oxygen transport and for other important metabolic reactions. Thus, fine regulation of this element is required since departures from its optimal levels produce severe alterations. For example, iron overload triggers the Fenton reaction in which the generation of hydroxyl radicals, causes mayor tissue damages [1]. By contrast, iron deficiency may cause anaemia. As in mammals there are no active mechanisms to eliminate iron, its absorption and storage regulation are crucial for maintaining appropriate levels of this mineral.

Among iron related diseases, anaemia is the most wide world spread, being considered as a pandemia. According to World Health Organization (WHO), a quarter of the world population suffers from anaemia, being women at fertile age the most affected group [2]. The prevalence of iron deficiency has been estimated to be between 8% and 33% of young women in Europe, 10%–16% in the United States, and 42% in developing countries [2].

Not only traditional dietary and host-related factors are determinants of iron levels, but also genetic factors play an important role. Mutations in key proteins involved in iron metabolism, red cell stability, and in iron absorption, have proved to generate severe anaemia and haemochromatosis [3-8]. Haemoglobin, serum ferritin, transferrin saturation with iron, which are parameters employed to assess iron status, show significant levels of heritability, and thus, seem to be genetically regulated [9,10]. In this context, recent studies have demonstrated that common allelic variants of certain genes are associated with iron related phenotypes [11,12], or account for a remarkably percentage of the genetic variation in the levels of iron-related parameters [13,14]. A genome wide linkage study revealed two regions that were related to the levels of transferrin receptor [15]: the higher was located on Chromosome 2 (lod score: 2.64) and the lower on Chromosome 3 (lod score: 1.94).

Therefore, by genotyping genetic determinants in a risk population for anaemia this study was aimed at two objectives: First, confirm in an independent sample, the importance of a specific group of genetic markers that has been already reported to be associated with several parameters related to iron status. Second, to follow up our previously observed linkage signals from a genome-wide linkage study [15] by testing three SNPs selected from the two regions (chromosome 2 and 3) that showed the before mentioned signals. In summary, the main goal was to identify common genetic variants associated with iron status especially with anaemia.

2. Material and methods

2.1 Subjects

A total of 284 subjects were recruited as part of a wide project aimed at understanding the interactions among iron, genes, nutrition and disease in menstruating women. As described in [16], the subjects were Caucasian women born in Europe, between 18 and 45 years of age, menstruating, non-pregnant, non-lactating and non-smoking. In this sample, 43.7% of the total individuals had normal haemoglobin (Hb) and ferritin (Ft) levels ($Hb \geq 12$, $Ft \geq 20$), 45.8% were iron deficient ($Hb \geq 12$ and $Ft < 20$ or $Hb \leq 12$ and $Ft > 20$), and 10.6% were iron deficient anaemic ($Hb < 12$ and $Ft < 20$). All women underwent a health questionnaire and blood test. The study was approved by the ethics committees of the *Clínica Puerta de Hierro* Hospital and the Spanish National Research Council, Madrid, Spain.

2.2 Hematological and Biochemical Determinations

Venous fasting blood samples (5 mL) were collected in EDTA tubes to analyse haemoglobin (g/dl), haematocrit (%), and mean corpuscular volume (MCV)(fl.) and in tubes with Gel+Clot activator to determine serum iron ($\mu\text{g/dl}$) and serum ferritin (ng/ml) by automatic standardized methods.

2.3 Genotyping

DNA was extracted from whole blood using standard phenol-chloroform methodology with proteinase K. Genotyping of 10 SNPs was carried out by a minisequencing method (Table 1), described in [16]. Criteria for selecting the 10 SNPs for minisequencing were reported in the same study. Briefly, these criteria were to choose those SNPs which were reported to be associated with biochemical

Table 1. Descriptive analysis of the 10 studied SNPs

SNP	Allele frequencies		Genotype frequencies			p-value, X ² (HWE)	Location
rs4820268	0.54(A)	0.46(G)	0.28(AA)	0.51(AG)	0.21(GG)	0.871493	TMPRSS6
rs855791	0.56(C)	0.44(T)	0.32(CC)	0.49(CT)	0.19(TT)	0.515545	TMPRSS6
rs1799852	0.88(C)	0.12(T)	0.76(CC)	0.24(CT)	0(TT)	0.125421	TRANSFERRIN
rs2280673	0.62(C)	0.38(A)	0.38(CC)	0.49(CA)	0.13(AA)	0.607449	TRANSFERRIN
rs1800562	0.97(G)	0.03(A)	0.93(GG)	0.07(GA)	0(AA)	0.505265	HFE
rs3811647	0.67(G)	0.33(A)	0.44(GG)	0.47(GA)	0.1(AA)	0.059705	TRANSFERRIN
rs2673289	0.61(C)	0.39(T)	0.38(CC)	0.47(CT)	0.15(TT)	0.906312	CHROMOSOME 2
rs1375515	0.67(A)	0.33(G)	0.44(AA)	0.45(AG)	0.1(GG)	0.655416	CACNA2D3
rs1799945	0.82(C)	0.18(G)	0.67(CC)	0.29(CG)	0.04(GG)	0.777062	HFE
rs16826756	0.81(A)	0.19(G)	0.65(AA)	0.31(AG)	0.04(GG)	0.960139	CHROMOSOME 2

parameters related to iron metabolism. A group of SNPs (rs3811647, rs1799852, rs2280673, rs1800562, rs855791 and rs1799945) were selected due to their association with haemoglobin levels and other iron-related parameters [17,18,19,20]. Another SNP (rs4820268) was selected because of its association with serum iron [13]. Finally three more SNPs were chosen from two linkage signals found in the Remacha et al. 2006 study [15]. These three tagSNPs were selected, by means of the The HapMap Project (<http://www.hapmap.org/>)[21]:

two from the linkage signal located on Chromosome 2 (rs16826756 and rs2673289), and another one from Chromosome 3 (rs1375515).

2.4 Statistical analysis

SNP allele frequencies assessment, genotype frequencies and Hardy-Weinberg equilibrium (HWE) were carried out. Normal distribution of hematological and biochemical parameters was assessed. All analysis was carried out using the R package (<http://www.Rproject.org/>) [22]. Serum ferritin was log transformed before it was analyzed, because it showed a skewed distribution.

The association between genotype and phenotype variables was studied by decision trees using SPSS vs.17 statistic package, and Bayesian Model Averaging (BMA) and linear regression models using R statistic package.

A decision tree was tested for each haematological and biochemical parameter, as well as for the distribution of the three clinical phenotypic groups (normal, iron deficient and anaemic) considering the latter as dependent variables and the 10 SNPs as independent variables. The decision trees are built by means of a non-parametric method that clusters the observations according to a factor or predictor that better explains the differences in the studied variable. Each generated subdivision is again partitioned according to the existence of new predictors with a significant effect, thus, the data are presented in a hierarchical manner [23]. This method extracts information by finding effects of the factors which are not homogeneous at all the levels of the dependent variable, and by discovering specific interactions among variables or data mining [24,25]. From all the different analyses available for this partitioning approach, those algorithms that apply non-binary divisions are more convenient. In our study, the CHAID (Chi-square Automatic Interaction Detection) exhaustive method was selected [26]. The significant threshold was $p < 0.05$, the minimum size of a node that could be divided consisted of 20 individuals and the minimum size of a child node consisted of 5 individuals. Finally this method applied corrections due to multiple testing through the Bonferroni method.

An analysis of the association between each of the biochemical parameters and the 10 SNPs through a BMA was carried out. BMA test accounts for the model uncertainty in the variable selection by averaging over the best models according to a posterior probability model. BMA test estimates all of the possible models and calculates a probabilistic average of the effect of each SNP.

In our study, using 10 SNPs, we considered $2^{10} = 1024$ possible models. The BMA posterior distribution of the effect of each SNP is a weighted average of its posterior distribution under all of the considered models [27].

Because results obtained with BMA and decision trees varied, linear regression models were calculated for the rs1375515 with all the biochemical parameters studied, testing for recessive and dominance models.

HapMap project (release #27) [21] was employed to describe the genomic location of rs1375515. Haploview 4.1 used the information of a 100kb region, extracted from the HapMap, to describe in detail the linkage disequilibrium patterns of the surroundings of rs1375515. Haplotype blocks were defined using a confident interval method describe elsewhere [28].

3. Results

Table 1 lists the allele and genotype frequencies, as well as the p-values for HWE test for the 10 studied SNPs. The 10 SNPs presented a Minimum Allele Frequency (MAF) greater than 0.01, and were in HWE.

3.1 BMA Test

Table 2 shows the results of the BMA test for MCV. The rs1375515 showed a significant association with MCV yielding a 55.4% probability value of being associated with this parameter, and appeared in the model which presented the highest posterior probability (0.225). The rest of the BMA tests did not include any of the SNPs in the best model according to its posterior probability and thus they are not presented.

Table 2. Association analysis between the 10 SNPs and Mean Corpuscular Volume (MCV) using BMA test

	p!=0	EV	SD	Model1	Model2	Model3	Model4	Model5
Intercept	100	87.7448111	0.60796	87.9715	87.242	88.3481	87.6379	88.3896
rs4820268	1.3	0.0002043	0.04914
rs855791	3.4	0.0116184	0.09927
rs1799852	7	0.0629203	0.29384
rs2280673	3.2	0.0099255	0.09574
rs1800562	7.1	0.1100278	0.5081
rs3811647	11.1	-0.0787761	0.2702	-0.6728
rs2673289	5	-0.0251253	0.14551
rs1375515	55.4	-0.6066949	0.63944	-1.1079	.	-1.0922	.	-1.07
rs1799945	5.1	0.0328943	0.18588
rs16826756	21.5	-0.2214793	0.49003	.	.	-1.0174	-1.0412	.
nVar				1	0	2	1	2
r2				0.021	0	0.034	0.013	0.029
BIC				-0.3901	0	1.5651	1.8608	3.0805
postprob				0.225	0.186	0.085	0.073	0.04

p!=0: Probability of the SNP to be associated with the variable. EV: Posterior mean of the beta parameter. SD: Standard deviation of each beta. Model1...5: Most probable multiple-SNP models. Post prob: Posterior probability of the model. nVar: Number of SNPs included in the model. BIC: Bayesian Information Criterion. r2: Coefficient of determination

3.2 Decision Trees

Only those decision trees that showed at least one significant subdivision (p-value<0.05) are described.

Figure 1 depicts the decision tree for haemoglobin. It is rs1375515 which better explains this variable, partitioning the sample ($P=0.019$) into two subgroups. Those individuals genotyped as A/A and those as A/G were clustered

together with a mean haemoglobin level of 13.189 g/dl. The rest of the sample, those typed as G/G homozygous were included in a different node with a mean value of 12.641 g/dl. This suggests that the absence of the A variant reduces the levels of haemoglobin by approximately 0.54 g/dl. As for the first node, it was split into two other subgroups ($P=0.045$), the rs1800562 (corresponding to the C282Y non-synonymous amino acid change) being the one producing this division. These two subgroups correspond to those individuals homozygous G/G and those heterozygous G/A. Thus, individuals carrying the allele A tend to express higher levels of haemoglobin (13.628 g/dl vs. 13.156 g/dl).

The tree built for \log_{10} ferritin is shown in Figure 2. It seems very similar to that of the haemoglobin, thus, the subdivisions have occurred in the same manner being the rs1375515 the one which produces the first split ($P=0.048$), and rs1800562 the one that produces the second ($P=0.038$). In the first subdivision, individuals A/A and A/G were clustered together yielding a mean of 1.284 (Ft values: 25.894 ng/ml) whereas individuals G/G presented a mean of 1.112 (Ft values: 18.033 ng/ml). The second division occurs in the group with higher mean, as in the previous tree. Considering this is a hierarchical method, the rs1375515, accounts for the variability of this parameter better than the rest of the SNPs, but with a lower significance level ($P=0.048$) in comparison with the haemoglobin tree.

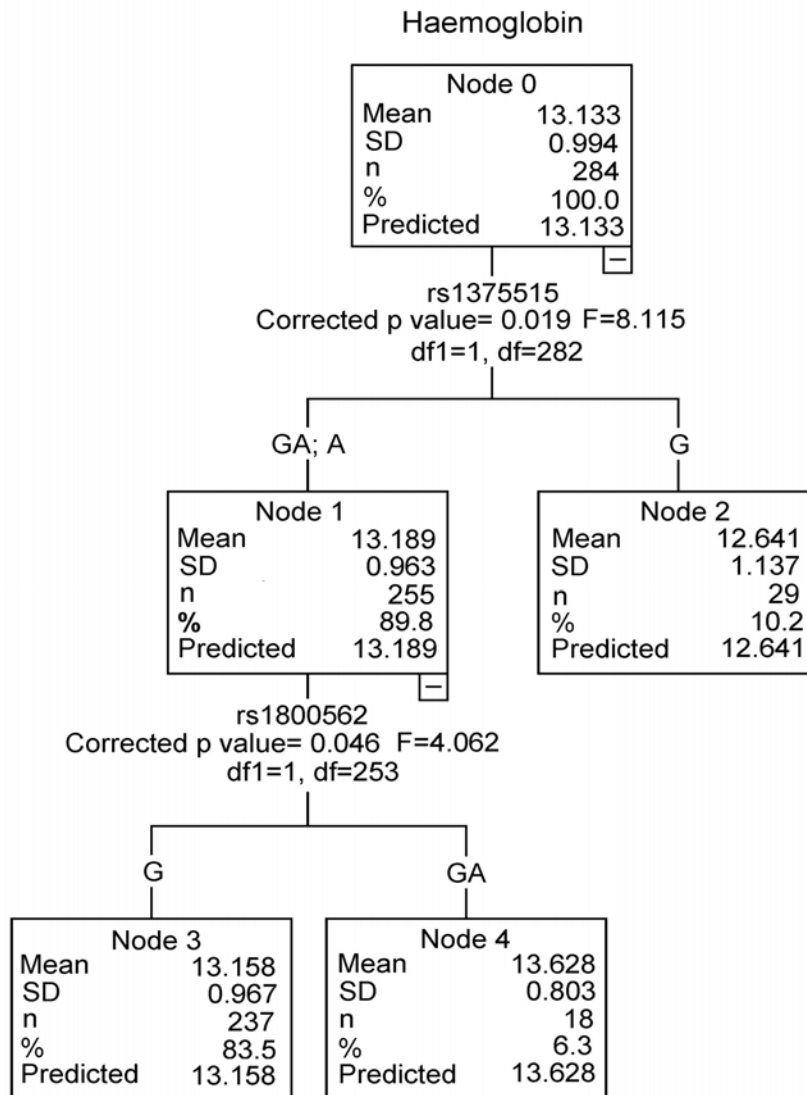


Fig. 1. Decision tree built with Haemoglobin (g/dl) as dependent variable and the 10 SNPs as independent variables or factors. %: Percentage of the total sample included in each node. Predicted: Predicted mean value assuming no effect for the factors

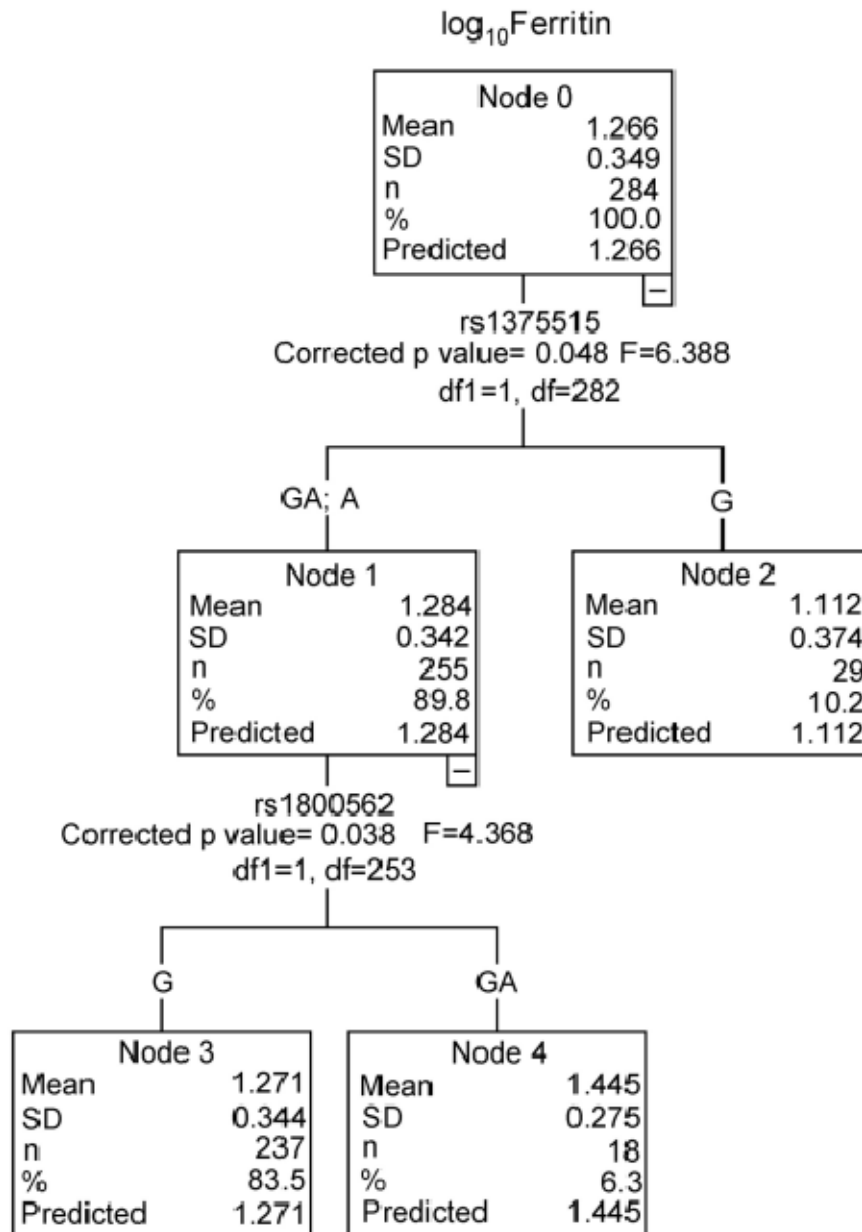


Fig. 2. Decision tree built with \log_{10} Ferritin (ng/ml) as dependent variable and the 10 SNPs as independent variables or factors. %: Percentage of the total sample included in each node. Predicted: Predicted mean value assuming no effect for the factors

Iron-clinical phenotypes

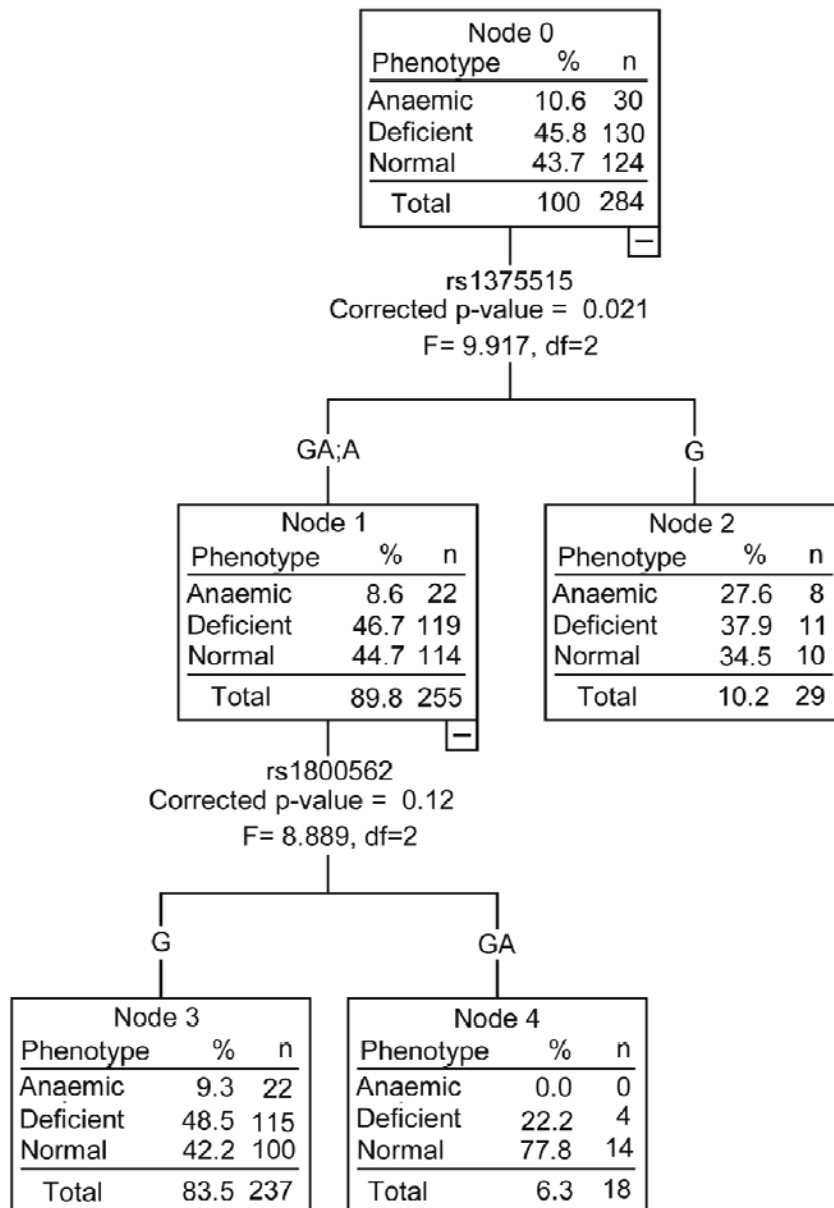


Fig. 3. Decision tree built with the distribution of iron clinical phenotypes (Iron deficient anaemic, iron deficient, and normal) as dependent categorical variable and the 10 SNPs as independent variables or factors. %: Percentage of each phenotype included in each node

Figure 3 shows the decision tree that assumes the distribution of the individuals in three phenotypic groups (anaemic, deficient and normal) as the dependent variable. As in the previous trees the first subdivision is caused by the

SNP rs1375515 generating two nodes ($P=0.02$). In node 2, which includes all the individuals homozygous for G/G, the distribution of phenotypes is significantly different from that of the whole sample (node 0), where the anaemic group having a sharp increase (27.6% vs. 10.6%) and both the normal and the deficient groups a considerable downscale. In node 2, only one out of 29 individuals presents the A mutation for rs1800562. Node 1 is subdivided into two groups by the SNP rs1800562. The resulting node 4, which contains those individuals heterozygous G/A, has a significantly different distribution of phenotypes, where there were no anaemic individuals (0%) and the proportion of normal individuals was highly increased compared to that of the total sample (77.8% in node 4 vs. 43.7% in the complete sample). This tree shows the significant combined effect of rs1375515 and rs1800562 over iron status, thus, in node 2, the effect of being homozygous G/G (rs1375515) and not carrying the A mutation of rs1800562 raises significantly the risk of anaemia, while in node 4 the combination of not being homozygous for the G mutation of rs1375515 and presenting the A mutation in rs1800562 produces a very low probability of anaemia and therefore increases the probability of being included in the normal group.

3.3 Dominance and recessive models

According to the decision trees, the association of the rs1375515 with haemoglobin levels and to a lesser extent with \log_{10} ferritin levels, raised questions about why this SNP did not appear significantly associated in the BMA test, for the same parameters.

The pattern the rs1375515 shows in the trees when producing the subdivisions, suggests that this SNP does not follow a co-dominance effect model. Thus, homozygous A/A and heterozygous A/G cluster in the same node, whereas G/G homozygous express a significant lower mean. This suggests that a recessive model may apply to these data. As BMA test assumes a co-dominant model, this could be the cause of the discrepancy between the two statistics.

Therefore, dominant and recessive models were tested for the rs1375515 with all the variables. The results presented in Table 3 show that this SNP generates

significant p-values for regression models considering recessive models, for the variables haemoglobin, \log_{10} ferritin, haematocrit and MCV.

Table 3. Simple linear regression analysis of variables versus rs1375515 assuming a recessive model

	Effect	Std. Error	t value	P-value	Variable
(Intercept)	13.1894	0.0615	214.55	0	Haemoglobin
rs1375515	-0.548	0.1924	-2.85	0.0047	
(Intercept)	2.9555	0.0499	59.28	0	\log_{10} Ferritin
rs1375515	-0.3943	0.156	-2.53	0.012	
(Intercept)	39.3071	0.169	232.56	0	Haematocrit
rs1375515	-1.2588	0.5289	-2.38	0.018	
(Intercept)	87.448	0.3121	280.19	0	MCV
rs1375515	-2.0169	0.9767	-2.07	0.0398	

3.4 HapMap and Haploview analysis

Figure 4 describes the genomic location of rs1375515, on chromosome 3 at position 54451680. This position maps in an intronic region of a gene that expresses an alpha voltage dependant calcium channel (*CACNA2D3*). The LD pattern of a surrounding region of 100kb is showed in Figure 4. According to Haploview rs1375515 is located in a 25kb haplotype block.

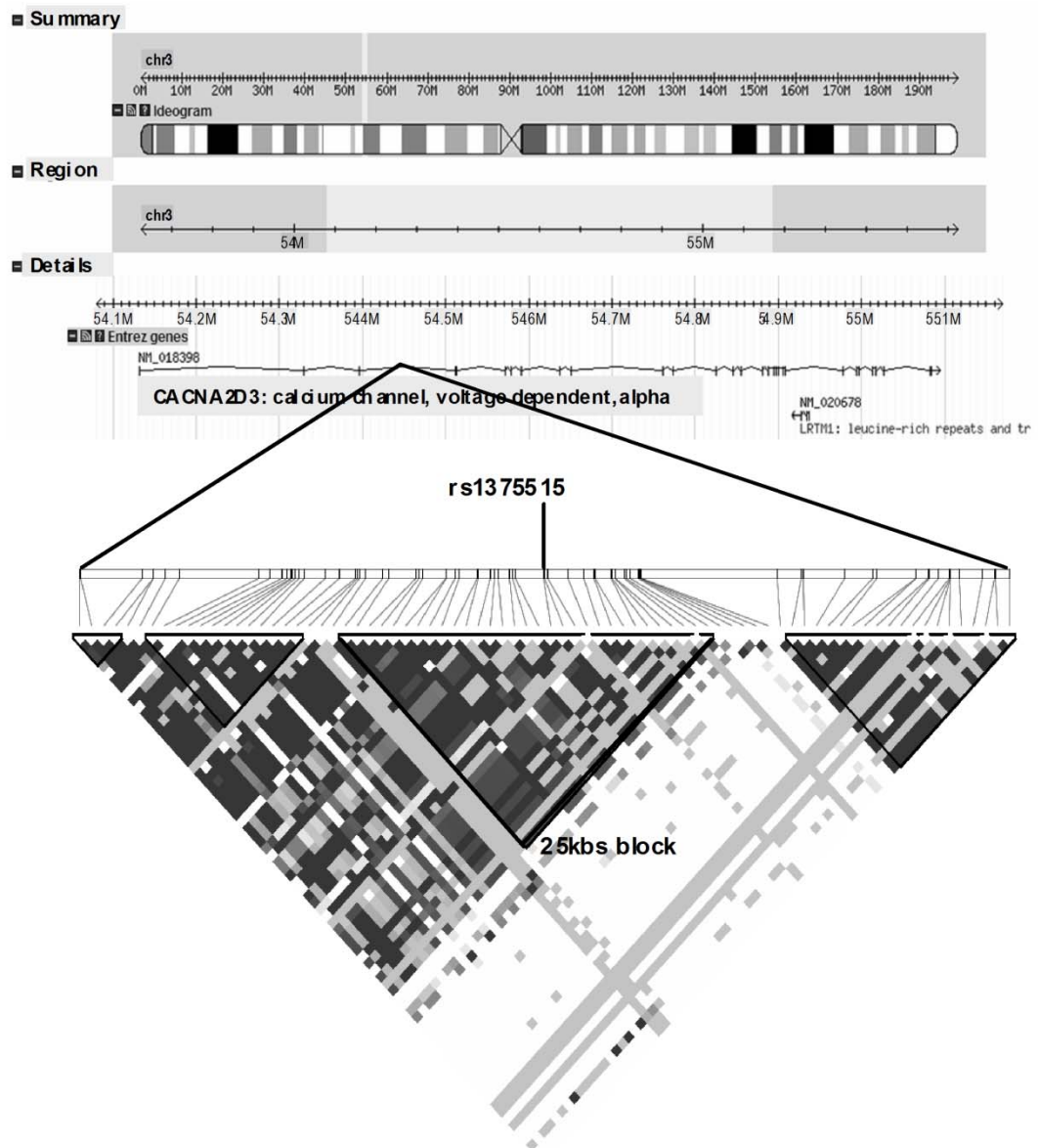


Fig. 4. Plots modified from HapMap (release # 27) and Haploview vs 4.1. HapMap plot shows a region of 1Mb in which the *CACNA2D3* is included. The linkage disequilibrium plot (values of r^2) was constructed from a 100kbs region in which rs1375515 is included. Darker diamonds indicate higher r^2 values

4. Discussion

The main goal of our study was to identify common genetic variants associated with general iron status. We show for the first time that rs1375515 is

associated with the levels of iron-related biomarkers, as well as with iron clinical phenotypes (normal, iron deficient and anaemic). According to BMA test, this SNP showed significant association with the levels of MCV. It also showed significant association with the levels of ferritin and haemoglobin according to decision trees. This is significant, since these two variables (haemoglobin and ferritin) are employed by the WHO to define healthy and pathological groups regarding iron status. The variant G in homozygosis was found in the current study to be significantly associated with higher risk of belonging to the anaemic group.

Although the two main statistics employed (BMA test and decision trees) did not present the same results, we should highlight that rs1375515 behaves in the same manner if we focus on its observed effect on iron status. Even though the results were different between the two statistical methods used (BMA and decision trees), the observed iron status was similar as regards of rs1375515. Regarding the decision trees, the allele G, or minor allele, is associated with low levels of both haemoglobin and ferritin which points to a lower iron status. Thus, for the BMA test for MCV, the coefficient of association in model1 shows that G allele is related to low levels of this parameter and it is known that low MCV together with low ferritin and haemoglobin values indicate iron deficiency anaemia risk. Therefore, from a physiological point of view, the presence of the G allele seems to be always associated with lower iron status.

Moreover, decision trees results suggest that rs1375515 had a recessive model effect on haemoglobin, ferritin, and on the clinical phenotypes. The results obtained using linear regression models were also consistent with the recessive effect hypothesis. Therefore, this could be the cause of not having found those associations in the BMA test given that this analysis only assumes codominant model effects.

Since rs1375515 is a tagSNP (according to HapMap criteria), it captures a great amount of the genetic variation in the region where it is placed. This SNP is located in the intronic region of a gene that expresses an alpha voltage dependant calcium channel (*CACNA2D3*). The significant associations that were found could

mean either that this SNP itself affects the parameters levels, or that there is another SNP in linkage disequilibrium (LD) with this one, which is responsible for the statistical association. The calcium channel gene, in which this SNP is located, spans approximately 1Mbs on chromosome 3, with rs1375515 separated approximately 750kbs from the 5' end and 300kbs from the 3' end of this gene. Linkage disequilibrium values in the genome can show different patterns among populations, and decrease with genetic distance, thus LD blocks greater than 100kbs are rare [29]. In this regard, rs1375515 is located, according to Haploview analysis, in a 25kb haplotype block which suggests that if there is another SNP, in LD with rs1375515 responsible for the statistical associations, it would be located most probably in the same block or in adjacent blocks within the calcium channel gene. Therefore, the data are pointing to a possible relation between variants of *CACNA2D3* gene and several iron status biomarkers (MCV, haemoglobin, ferritin) as well as with iron clinical phenotypes.

Regarding the association with MCV, given the position of rs1375515 in a calcium gene, there are several evidences in the bibliography that relate cell volume and ionic channels with iron metabolism pathologies that should be discussed as follows. Fine regulation of the cell volume, associated with ionic exchange [30,31], is essential for normal cellular function, and consequently, alterations in cell volume participate in the physiopathology of disorders such as liver insufficiency, fibrosing disease, and sickle anaemia [33]. Moreover, red cells are extremely sensitive to volume changes that may easily produce haemolytic related anaemias [33]. Recently, Sebastiani et al. [34] have discovered some genetic modifiers of the severity of sickle cell anaemia, some of them within a K⁺ channel *KCNK6*. Although this association is not fully understood, this work may support the fact that cationic exchange is somehow associated with erythrocytes volume and therefore with deficiency anaemia or other anaemias.

According to the bibliography, calcium channels are reported, to be associated with cell volume and iron status as well. In Lew et al. [33] it was shown how blocking these channels affected red cell volume by dehydration. This resulted,

as well, in changes in cell volume and anaemia. In the same line, it has been demonstrated [35] that blocking calcium channels had beneficial effects on an iron overload cardiomyopathy. Other works show that these channels could be directly responsible for iron uptake, namely the L-type voltage dependent calcium channel could be an alternate route for iron to enter the cardiomyocytes [36].

All these evidences may suggest that the G variant of the rs1375515, or other variants in LD, can be involved in hampering the calcium channel function which could increase the risk of anaemia. The intronic location of this SNP should not be an argument to discard its possible effect, as other SNPs located in intronic regions have been proved to influence the levels of the protein expression [13,34].

We found that the C282Y substitution (rs1800562) is related to high levels of haemoglobin and ferritin, and to low probabilities of being anaemic, counterbalancing to some extent the effect of the G allele of rs1375515. In fact, the results have shown the combined effect of these two SNPs can significantly vary the distribution of the iron clinical phenotypes. This SNP (rs1800562) located at the *HFE* gene, is associated with haemochromatosis [19,37,38] and has been reported to show significant effect on haemoglobin levels as well as other parameters related to iron status such as serum ferritin, serum iron, serum transferrin and transferrin saturation [13,17]. This study supports other recent findings in which common variants the *HFE* gene [18], may be used as predictor factors of iron status.

The rs855791 and the rs4820268, both located in *TMPRSS6* gene, were analyzed because they have recently been associated with haemoglobin and iron status, MCV [13,18,39], serum iron [17] and to some iron-related diseases such as iron refractory anaemia [5]. In our study, these SNPs had no significant effect on the levels of any of the parameters regardless of the statistic employed (data not shown). A possible explanation could be that the relatively small size of our sample (n=284) hampers to find associations for SNPs that do not show very strong effect. In this respect the effect of the recessive homozygous of rs855791 on haemoglobin levels in Chambers et al. 2009 [18] was a 0.21 g/dl reduction (in a larger sample),

whereas in the same case for rs1375515 (present study) the reduction was of 0.54 g/dl, which is notably higher.

The region in which rs2673289 and rs16826756 are located (Chromosome 2), showed a linkage signal with transferrin receptor in a previous family-based study [15], however, in the present work these SNPs have not showed any association with the studied parameters. An explanation for this could be that the markers that produced the linkage signal, may be rare or specific of some families, and hence, are not present in our sample.

It is important to note that our sample was not chosen to be representative of the general population. The individuals were chosen to be women at fertile age since this is the population group most at risk of suffering from anaemia. Following the WHO criteria for haemoglobin and ferritin levels, in our sample the proportion of those defined as iron deficient and anaemic (45.8% and 10.6%) compared with those defined as healthy (43.7%) could be overrepresented compared with the general population.

In summary, our results show that the rs1375515 is associated with haematological and biochemical parameters used to assess iron status. The location of this SNP in a calcium channel gene (*CACNA2D3*), suggests that the functionality of this channel regarding its relation to iron related parameters and anaemia should be further investigated. The combined effect of this SNP together with the C282Y substitution (rs1800562) explains significant differences in the risk of developing anaemia. Thus, our study has broad implications for future studies that focus on the basis of genetic variation within iron-related traits.

Acknowledgments

This study was supported partially by projects AGL2009-11437, N8/2006-4130063, PI-08/0420 and PI-08/0756. R Blanco-Rojo was supported by a JAE-predoc grant from CSIC and Social European Found, S Bertoncini was supported by Grupo Santander 2009 (Estancia doctores y tecnólogos UCM, and J.M. Soria by

“Programa d'Estabilització d'Investigadors de la Direcció d'Estratègia i Coordinació del Departament de Salut” (Generalitat de Catalunya).

References

- [1] M.W. Hentze, M.U. Muckenthaler and N.C. Andrews, Balancing acts: molecular control of mammalian iron metabolism, *Cell* **117** (2004), 285-297.
- [2] B. de Benoist, E. McLean, I. Egli and M. Cogswell, Worldwide prevalence of anaemia 1993-2005: WHO Global Database on Anaemia. Geneva, World Health Organization, 2008.
- [3] K.E. Finberg, Iron-refractory iron deficiency anemia, *Semin Hematol* **46** (2009), 378-386.
- [4] S.G. Gehrke, A. Pietrangelo, M. Kascák, A. Braner, M. Eisold, H. Kulaksiz, T. Herrmann, U. Hebling, K. Bents, R. Gugler and W. Stremmel, HJV gene mutations in European patients with juvenile hemochromatosis, *67 Clin Genet* (2005), 425-428.
- [5] A.J. Ramsay, V. Quesada, M. Sanchez, C. Garabaya, M.P. Sardà, M. Baiget, A. Remacha, G. Velasco and C. López-Otín, Matriptase-2 mutations in iron-refractory iron deficiency anemia patients provide new insights into protease activation mechanisms, *Hum Mol Genet* **18** (2009), 3673-3683.
- [6] A. Roetto, F. Daraio, F. Alberti, P. Porporato, A. Calì, M. De Gobbi and C. Camaschella, Hemochromatosis due to mutations in transferrin receptor 2, *Blood Cells Mol Dis* **29** (2002), 465-470.
- [7] B. Sarria, S. Navas-Carretero, A.M. Lopez-Parra, A.M. Pérez-Granados, E. Arroyo-Pardo, M.A. Roe, B. Teucher, M.P. Vaquero and S.J. Fairweather-Tait, The G277S transferrin mutation does not affect iron absorption in iron deficient women, *Eur J Nutr* **46** (2007), 57-60.
- [8] D.J. Weatherall. Pathophysiology of thalassaemia, *Baillieres Clin Haematol* **11** (1998), 127-146.
- [9] O.T. Njajou, B.Z. Alizadeh, Y. Aulchenko, M.C. Zillikens, H.A. Pols, B.A. Oostra, D.W. Swinkels and C.M. van Duijn, Heritability of serum iron, ferritin and transferrin saturation in a genetically isolated population, the Erasmus Rucphen Family (ERF) Study, *Hum Hered* **61** (2006), 222-228.
- [10] J.B. Whitfield, L.M. Cullen, E.C. Jazwinska, L.W. Powell, A.C. Heath, G. Zhu, D.L. Duffy and N.G. Martin, Effects of HFE C282Y and H63D polymorphisms and polygenic background on iron stores in a large community sample of twins, *Am J Hum Genet* **66** (2000), 1246-1258.
- [11] H.K. Bayele and S.K. Srail, Genetic variation in hepcidin expression and its implications for phenotypic differences in iron metabolism, *Haematologica* **94** (2009), 1185-1188.

- [12] C.E. McLaren, C.P. Garner, C.C. Constantine, S. McLachlan, C.D. Vulpe, B.M. Snively, V.R. Gordeuk, D.A. Nickerson, J.D. Cook, C. Leiendecker-Foster, K.B. Beckman, J.H. Eckfeldt, L.F. Barcellos, J.A. Murray, P.C. Adams, R.T. Acton, A.A. Killeen and G.D. McLaren, Genome-Wide Association Study Identifies Genetic Loci Associated with Iron Deficiency. *Plos One* (2011), 6: e17390. doi:10.1371/journal.pone.0017390
- [13] B. Benyamin, A.F. McRae, G. Zhu, S. Gordon, A.K. Henders, A. Palotie, L. Peltonen, N.G. Martin, G.W. Montgomery, J.B. Whitfield and P.M. Visscher, Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels, *Am J Hum Genet* **84** (2009), 60-65.
- [14] N. Soranzo, T.D. Spector, M. Mangino et al., A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium, *Nat Genet* **41** (2009), 1182-1190.
- [15] A.F. Remacha, J.C. Souto, J.M. Soria, A. Buil, M.P. Sardà, M. Lathrop, J. Blangero, L. Almasy and J. Fontcuberta, Genomewide linkage analysis of soluble transferrin receptor plasma levels, *Ann Hematol* **85** (2006), 25-28.
- [16] S. Bertoncini, R. Blanco-Rojo, C. Baeza, E. Arroyo-Pardo, M.P. Vaquero and A.M. López-Parra, A novel SNaPshot assay to detect genetic mutations related to iron metabolism, *Genet Test Mol Biomark* **15** (2011), 173-179.
- [17] B. Benyamin, M.A. Ferreira, G. Willemsen, S. Gordon, R.P. Middelberg, B.P. McEvoy, J.J. Hottenga, A.K. Henders, M.J. Campbell, L. Wallace, I.H. Frazer, A.C. Heath, E.J. de Geus, D.R. Nyholt, P.M. Visscher, B.W. Penninx, D.I. Boomsma, N.G. Martin, G.W. Montgomery and J.B. Whitfield, Common variants in TMPRSS6 are associated with iron status and erythrocyte volume, *Nat Genet* **41** (2009), 1173-1175.
- [18] J.C. Chambers, W. Zhang, Y. Li, J. Sehmi, M.N. Wass, D. Zabaneh, C. Hoggart, H. Bayele, M.I. McCarthy, L. Peltonen, N.B. Freimer, S.K. Sriv, P.H. Maxwell, M.J. Sternberg, A. Ruukonen, G. Abecasis, M.R. Jarvelin, J. Scott, P. Elliott and J.S. Kooner, Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels, *Nat Genet* **41** (2009), 1170-1172.
- [19] J.N. Feder, A. Gnirke, W. Thomas, et al., A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis, *Nat Genet* **13** (1996), 399-408.
- [20] S.K. Ganesh, N.A. Zakai, F.J. van Rooij, et al., Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium, *Nat Genet* **41** (2009), 1191-1198.
- [21] The International HapMap Project, *Nature* **426** (2003), 789-796.
- [22] R.D.C. Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.Rproject.org/>
- [23] E. Marvez, S.J. Weiss, D.E. Houry and A.A. Ernst, Predicting adverse outcomes in a diagnosis-based protocol system for rapid sequence intubation, *Am J Emerg Med* **21** (2003), 23-29.
- [24] D. Delen, G. Walker and A. Kadam, Predicting breast cancer survivability: a comparison of three data mining methods, *Artif Intell Med* **34** (2005), 113-127.

- [25] I.S. Ríos, C. Bielza and A. Mateos, Fundamentos de los Sistemas de Ayuda a la Decisión, ed., Ra- Ma, Madrid, España, 2001.
- [26] M.T. Martín, M.V. Román and J.P. Lévy. Criterios de diferenciación de segmentos de predisposición tecnológica en el ámbito hospitalario público, *Ciencia Ergo Sum* **12** (2005), 125-132.
- [27] A.E. Raftery, Bayesian model selection in social research, *Sociological Methodology* **25** (1995), 111-163.
- [28] S.B. Gabriel, S.F. Schaffner, H. Nguyen, J.M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, S.N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E.S. Lander, M.J. Daly and D. Altshuler, The structure of haplotype blocks in the human genome, *Science* **296** (2002), 2225-2229.
- [29] M.J. Daly, J.D. Rioux, S.F. Schaffner, T.J. Hudson and E.S. Lander, High-resolution haplotype structure in the human genome, *Nat Genet* **29** (2001), 229-232.
- [30] L. McManus, K.B. Churchwell and K. Strange, Regulation of cell volume in health and disease . *N Engl J Med* **333** (1995), 1260-1266.
- [31] T. Tiffert, N. Daw, Z. Etzion, R.M. Bookchin and V.L. Lew, Age decline in the activity of the Ca²⁺-sensitive K⁺ channel of human red blood cells, *J Gen Physiol* **129** (2007), 429-436.
- [32] F. Lang, Mechanisms and significance of cell volume regulation, *J Am Coll Nutr* **26** (2007), 613S-623S.
- [33] V.L. Lew, Z. Etzion and R.M. Bookchin, Dehydration response of sickle cells to sickling-induced Ca(++) permeabilization. *Blood* **99** (2002), 2578-2585.
- [34] P. Sebastiani, N. Solovieff, S.W. Hartley, J.N. Milton, A. Riva, D.A. Dworkis, E. Melista, E.S. Klings, M.E. Garrett, M.J. Telen, A. Ashley-Koch, C.T. Baldwin and M.H. Steinberg, Genetic modifiers of the severity of sickle cell anemia identified through a genome-wide association study, *Am J Hematol* **85** (2010), 29-35.
- [35] K. Sugishita, M. Asakawa, S. Usui and T. Takahashi, A case of iron overload cardiomyopathy: beneficial effects of iron chelating agent and calcium channel blocker on left ventricular dysfunction, *Int Heart J* **50** (2009), 829-838.
- [36] J.A. Gaasch, W.J. Geldenhuys, P.R. Lockman, D.D. Allen and C.J. Van der Schyf, Voltage-gated Calcium Channels Provide an Alternate Route for Iron Uptake in Neuronal Cell Cultures, *Neurochem Res* **32** (2007), 1686-1693.
- [37] A.F. Remacha, M.P. Sarda, M.J. Barcelo et al., Genotype-phenotype correlation in a Spanish population homozygous for the H63D mutation of the HFE gene, *Ann Hematol* **85** (2006), 340-342.
- [38] M. Sánchez, M. Bruguera, J. Bosch, J. Rodes, F. Ballesta and R. Oliva, Prevalence of the Cys282Tyr and His63Asp HFE gene mutations in Spanish patients with hereditary hemochromatosis and in controls, *J Hepatol* **29** (1998), 725-728.
- [39] X. Du, E. She, T. Gelbart et al., The serine protease TMPRSS6 is required to sense iron deficiency, *Science* **320** (2008), 1088-1092.

4.2. “Genetic contribution to iron status: SNPs related to iron deficiency anaemia and fine mapping of CACNA2D3 calcium channel subunit”

2015

Blood Cells, Molecules and Diseases. 2015;55(4):273-80.

doi:10.1016/j.bcmed.2015.07.008

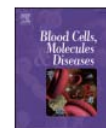
ISSN 1079-9796

Blood Cells, Molecules and Diseases 55 (2015) 273–280



Contents lists available at ScienceDirect

Blood Cells, Molecules and Diseases

journal homepage: www.elsevier.com/locate/bcmd

Genetic contribution to iron status: SNPs related to iron deficiency anaemia and fine mapping of CACNA2D3 calcium channel subunit



Carlos Baeza-Richer ^{a,*}, Eduardo Arroyo-Pardo ^a, Ruth Blanco-Rojo ^b, Laura Toxqui ^c, Angel Remacha ^d, M. Pilar Vaquero ^c, Ana M. López-Parra ^a

^a Department of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Madrid, Spain

^b Lipids and Atherosclerosis Unit, Reina Sofia University Hospital, Maimonides Institute for Biomedical Research at Cordoba (IMIBIC), University of Cordoba, Spain

^c Department of Metabolism and Nutrition, Institute of Food Science, Technology and Nutrition (ICTAN-CSIC), Madrid, Spain

^d Haematology Service, Laboratory and Analyses, Sant Pau Hospital, Barcelona, Spain

ARTICLE INFO

Article history:
Submitted 8 June 2015
Revised 13 July 2015
Accepted 14 July 2015
Available online 16 July 2015

Keywords:
Iron deficiency anaemia
SNPs
Association study
Calcium channel
Iron uptake

ABSTRACT

Numerous studies associate genetic markers with iron- and erythrocyte-related parameters, but few relate them to iron-clinical phenotypes. Novel SNP rs1375515, located in a subunit of the calcium channel gene *CACNA2D3*, is associated with a higher risk of anaemia. The aim of this study is to further investigate the association of this SNP with iron-related parameters and iron-clinical phenotypes, and to explore the potential role of calcium channel subunit region in iron regulation. Furthermore, we aim to replicate the association of other SNPs reported previously in our population. We tested 45 SNPs selected via systematic review and fine mapping of *CACNA2D3* region, with haematological and biochemical traits in 358 women of reproductive age. Multivariate analyses include back-step logistic regression and decision trees. The results replicate the association of SNPs with iron-related traits, and also confirm the protective effect of both A allele of rs1800562 (*HFE*) and G allele of rs4895441 (*HBS1L-MYB*). The risk of developing anaemia is increased in reproductive age women carriers of A allele of rs1868505 (*CACNA2D3*) and/or T allele of rs13194491 (*HIST1H2BJ*). Association of SNPs from fine mapping with ferritin and serum iron suggests that calcium channels could be a potential pathway for iron uptake in physiological conditions.

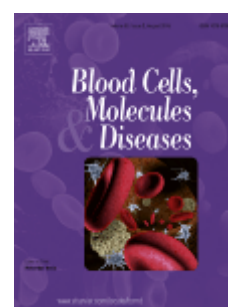
© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Iron homeostasis is essential for numerous physiological processes. However, excess of iron could form toxic free radicals, and it thus requires precise regulation [1]. Because humans do not possess an active mechanism for iron excretion, this regulation is carried out mainly by modulating the uptake of iron from the diet by enterocytes and transfer of this iron to the systemic circulation [2]. The key proteins that control this process also control the release of stored iron to plasma, to supply iron for erythropoiesis and other metabolic activities [3]. Small deviations from correct iron levels are the ultimate cause of several disorders, of which iron deficiency anaemia (IDA) is the most prevalent. Indeed, it is considered a pandemic according to the WHO, which classifies women in fertile age as the most at-risk group [4]. Moreover, iron

biochemical parameters used to assess iron status, such as haemoglobin, haematocrit, ferritin, mean cell volume (MCV), transferrin and serum iron, supports this hypothesis [12–14]. Several studies on humans have associated many genetic markers, mainly SNPs, with iron-related parameters [15–19] and with anaemia itself [20–22]. Despite the vast amount of information that these genome-wide association studies (GWAS) have made available, it is still difficult to replicate or apply the obtained results in practical health contexts [23–25]. In this respect, the increasing complexities of health and illness issues render it necessary to broaden methodological strategies in research, and the use of data mining methods has been suggested as a means to integrate all of these associations [26].

Although some of the SNPs associated with iron metabolism parameters are found in protein coding regions, most of them, as the Encode



4.2 “Genetic contribution to iron status: SNPs related to iron deficiency anaemia and fine mapping of CACNA2D3 calcium channel subunit” Artículo

Carlos Baeza-Richer^a, Eduardo Arroyo-Pardo^a, Ruth Blanco-Rojo^b, Laura Toxqui^c, Angel Remacha^d, M. Pilar Vaquero^c, Ana M. López-Parra^a

^a Department of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Madrid, Spain

^b Lipids and Atherosclerosis Unit, Reina Sofia University Hospital, Maimonides Institute for Biomedical Research at Cordoba (IMIBIC), University of Cordoba

^c Department of Metabolism and Nutrition, Institute of Food Science, Technology and Nutrition (ICTAN-CSIC), Madrid, Spain

^d Haematology Service, Laboratory and Analyses. Sant Pau Hospital, Barcelona

Author for correspondence: Carlos Baeza Richer, e-mail: cbaezaricher@med.ucm.es

Resumen

Varios trabajos relacionan regiones genéticas con parámetros asociados al hierro y a los eritrocitos, pero son menos los que los relacionan con los fenotipos clínicos de la anemia ferropénica. El SNP rs1375515, localizado en una subunidad de un canal de calcio, se ha asociado con mayor riesgo de sufrir anemia. El objeto de este estudio es investigar en profundidad la relación de este SNP con los parámetros del hierro y los fenotipos clínicos, y explorar las posibles implicaciones del canal del calcio en la regulación del hierro. Además, se busca replicar en nuestra población la asociación previamente descrita de otros SNPs con estos parámetros. Se testaron 45 SNPs seleccionados a través de una búsqueda sistemática de la literatura y de un *Fine Mapping* de la región *CACNA2D3*, con los rasgos hematológicos y bioquímicos de 358 mujeres en edad fértil. Se emplearon análisis multivariantes como regresión logística por pasos hacia atrás y árboles de decisión. Los resultados replican la asociación de SNPs relacionados anteriormente con parámetros del hierro y confirman el efecto protector del alelo A de rs1800562 (*HFE*) y del G de rs4895441 (*HBS1L-MYB*). El riesgo de sufrir anemia se ve incrementado en los portadores del alelo A de rs1868505 (*CACNA2D3*) y/o del alelo T de rs1319441 (*HIST1H2BJ*). La asociación encontrada entre los SNPs del

Fine Mapping con la ferritina y el hierro sérico sugiere que los canales de calcio podrían ser una vía directa para la entrada del hierro en condiciones fisiológicas.

Palabras clave: *anemia ferropénica, SNPs, estudios de asociación, canal de calcio, absorción de hierro*

Abstract

Numerous studies associate genetic markers with iron- and erythrocyte-related parameters, but few relate them to iron-clinical phenotypes. Novel SNP rs1375515, located in a subunit of the calcium channel gene *CACNA2D3*, is associated with a higher risk of anaemia. The aim of this study is to further investigate the association of this SNP with iron-related parameters and iron-clinical phenotypes, and to explore the potential role of calcium channel subunit region in iron regulation. Furthermore, we aim to replicate the association of other SNPs reported previously in our population. We tested 45 SNPs selected via systematic review and fine mapping of *CACNA2D3* region, with haematological and biochemical traits in 358 women of reproductive age. Multivariate analyses include back-step logistic regression and decision trees. The results replicate the association of SNPs with iron-related traits, and also confirm the protective effect of both A allele of rs1800562 (*HFE*) and G allele of rs4895441 (*HBS1L-MYB*). The risk of developing anaemia is increased in reproductive age women carriers of A allele of rs1868505 (*CACNA2D3*) and/or T allele of rs13194491 (*HIST1H2BJ*). Association of SNPs from fine mapping with ferritin and serum iron suggests that calcium channels could be a potential pathway for iron uptake in physiological conditions.

Keywords: *iron deficiency anaemia, SNPs, association study, calcium channel, iron uptake*

1. Introduction

Iron homeostasis is essential for numerous physiological processes. However, excess of iron could form toxic free radicals, and it thus requires precise regulation [1]. Because humans do not possess an active mechanism for iron excretion, this regulation is carried out mainly by modulating the uptake of iron from the diet by enterocytes and transfer of this iron to the systemic circulation [2]. The key proteins that control this process also control the release of stored iron to plasma, to supply iron for erythropoiesis and other metabolic activities [3]. Small deviations from correct iron levels are the ultimate cause of several disorders, of which iron deficiency anaemia (IDA) is the most prevalent. Indeed, it is considered a pandemic according to the WHO, which classifies women in fertile age as the most at-risk group [4]. Moreover, iron deficiency is associated with other disorders [5,6].

IDA is influenced by diet, iron intake, blood loss, physiological status and infections [7,8]. However, the evidence on several genetic disorders of iron metabolism in human and animals indicates a plausible genetic contribution to iron regulation [9-11]. Moreover, heritability of the biochemical parameters used to assess iron status, such as haemoglobin, haematocrit, ferritin, mean cell volume (MCV), transferrin and serum iron, supports this hypothesis [12-14]. Several studies on humans have associated many genetic markers, mainly SNPs, with iron-related parameters [15-19] and with anaemia itself [20-22]. Despite the vast amount of information that these genome-wide association studies (GWAS) have made available, it is still difficult to replicate or apply the obtained results in practical health contexts [23-25]. In this respect, the increasing complexities of health and illness issues render it necessary to broaden methodological strategies in research, and the use of data mining methods has been suggested as a means to integrate all of these associations [26].

Although some of the SNPs associated with iron metabolism parameters are found in protein codifying regions, most of them, as the Encode project is revealing [27,28], are not involved in structural changes but in functional aspects. The

authors of the present study found a novel SNP (rs1375515), located in an intronic region of a gene that codifies a subunit of a calcium channel *CACNA2D3*, significantly associated with haemoglobin and ferritin, as well as with different probabilities of belonging to different iron-clinical phenotypes [21]. Calcium channels are associated with iron homeostasis through the modification of red cell volume that leads to disorders such as liver insufficiency, fibrosing disease and sickle cell anaemia [29]. Other studies have shown that these channels could be directly responsible for iron uptake, namely that the L-type voltage dependent calcium channel could be an alternate route for iron to enter different cell types [30,31].

The aims of this study are to replicate the association of the most relevant SNPs not only with iron-related parameters but also with iron-clinical phenotypes, using data mining and qualitative methods, and to explore the potential role of the calcium channel subunit region by means of fine mapping.

2. Material and Methods

Subjects

A total of 358 women of reproductive age were recruited in a wider project aimed at understanding the interactions among iron, genes, nutrition and disease in menstruating women. This sample is an extension in 74 women of the sample analyzed in Bertoncini et al. and Baeza-Richer et al. [21,32]. The subjects were Caucasian women born in Europe, aged between 18 and 45 years old, menstruating, non-pregnant, non-lactating and non-smoking. In this sample, 43.6% of the total individuals had normal haemoglobin (Hb, g/dl) and ferritin (Ft, µg/l) levels ($Hb \geq 12$, $Ft \geq 20$), 45.8% were iron deficient ($Hb \geq 12$ and $Ft < 20$ or $Hb \leq 12$ and $Ft > 20$) and 10.6% were iron deficient anaemic ($Hb < 12$ and $Ft < 20$). This classification follows WHO criteria [4]. All women were administered a health questionnaire and underwent a blood test. The study was approved by the ethics

committee of the *Puerta de Hierro* Hospital and the Spanish National Research Council, Madrid, Spain.

SNP selection

Thirty five SNPs were analysed in this cohort for the first time, plus 10 SNPs already typed in Bertoni et al. and Baeza-Richer et al.[21,32] where these 10 SNPs were analysed in 284 out of the whole sample. The total 45 SNPs (Table 1S) were selected for this study according to the following criteria:

a) Fine Mapping: this technique allows us to look deeper into the possible implication of *CACNA2D3* in iron metabolism, by genotyping more extensively this region. In general terms, pairwise linkage disequilibrium (LD) gives you information of how two SNPs vary together, so that you need to select less SNPs to capture all the genetic variability (information) of a given region. Using Tagger [33] from the Haploview software [34] vs 4.1, we fine-mapped a 100kb LD (linkage disequilibrium) block region in which rs1375515 is located (*CACNA2D3*). Haploview software will select the least SNPs needed to capture most of the region variability employing LD analysis. The selection criteria were: pairwise tagging only, and a LD value (r^2) of 0.8. After selection, we eliminated those SNPs with a MAF (minimum allele frequency) lower than 0.01 from the analysis. The final selection included 16 SNPs.

b) Extensive and systematic literature search: 19 SNPs from (GWAS) were selected employing the following criteria: highly significant association with iron-related parameters in at least two independent studies, and associated with at least one of the parameters under analysis: haemoglobin, ferritin, haematocrit, MCV, serum iron and transferrin. SNPs not in LD were selected.

c) SNaPshot Assay: as described before [21,32], we selected another 10 SNPs which were included in the mini-sequencing assay developed by our group. The SNPs included in this assay were selected in our previous study because of their strong association with iron-related traits. They are located in the *TMPRSS6*, *HFE* and *TFR* genes. Three other SNPs (including rs1375515) were selected from a linkage study [35].

Genotyping

Ten SNPs were genotyped using a multiplex minisequencing assay (SNaPshot) that was developed by our group [32]. Genotypes were obtained in a 310 ABI Sequencer. The other 35 SNPs were genotyped by means of the Sequenom technique, using iPLEX chemistry. SNP genotyping services were provided by the Spanish "National Genotyping Centre" (CEGEN-USC, <http://www.cegen.org>).

Haematological and biochemical determinations

Fasting venous blood samples (5 ml) were collected in EDTA tubes to analyse haemoglobin (g/dl), haematocrit (%) and MCV (fl.), and in tubes with Gel + Clot activator to determine serum iron ($\mu\text{g/dl}$), serum ferritin (ng/ml) and serum transferrin (mg/dl) by automatic standardised methods.

Statistical analysis

SNP allele frequencies, genotype frequencies and the Hardy-Weinberg equilibrium were determined. Normal distribution of haematological and biochemical parameters was assessed. These analyses were carried out using the R package (<http://www.Rproject.org/>). Serum ferritin and MCV were log-transformed prior to analysis because they showed a skewed distribution.

The association between genotype and phenotype was studied using decision trees and linear regression models. Decision trees are one of the most popular approaches for representing classifiers [26]. This algorithm usually works top-down by choosing a variable at each step (in this case the SNPs) that is the most significant variable to split the set of items (the sample in this case). The best SNP is defined by how well it splits the set into homogeneous subsample that have the same value regarding the SNP genotypes. Each generated subsample could be again partitioned according to the existence of new SNPs with a significant effect, therefore the data are presented in a hierarchical way. Decision trees uses a Chi-Square test that eventually could detect interactions among the SNPs(CHAID:Chi-Square Automatic Interaction Detection). In order to avoid hazardous results a significance threshold of $p < 0.05$, a minimum parent node size of 30 individuals,

and a minimum child node size of 10 was used. Finally, this method applies Bonferroni corrections due to multiple testing. This analysis was performed using the statistical package SPSS vs.19.

In addition, a back-step logistic regression model was tested which assumed an association between iron-clinical phenotypes and the SNPs. This method is suitable for qualitative responses such as iron-clinical phenotypes (normal iron deficiency and IDA). As in general regressions, it assumes a linear response between explanatory and response variables [36]. In the first step, the back-step model analyses all the SNPs in the model and chooses the most significant one. In the following steps, it selects the most significant SNPs, maintaining those already chosen, until all significant SNPs are included. This model is appropriate to assess the degree of prediction of an individual presenting a certain iron status (iron-clinical phenotypes) regarding a specific combination of those SNPs selected. This analysis was carried out using SAS v.9.2.

3. Results

Association of SNPs with iron-related parameters

The regression analysis results for all the SNPs with the iron-related parameters, testing several effect models, are shown in Table 1. Some SNPs seemed to be specifically associated with one parameter, such as rs3811647 (*TF*) and rs198846 (*TF*) with transferrin, rs11970772 (*CCND3*) with serum iron, and rs16826756 (chr 2) with MCV, whereas others were associated with several parameters, such as rs4895441 (*HBS1L*), rs13194491 (*HIST1H2BJ*), rs1799852 (*TF*), and rs2673289 (chr 2). Parameters such as haemoglobin and haematocrit shared several SNPs, whereas transferrin showed different significant SNPs from all the other parameters. Fine-mapped SNPs were only observed to be significantly associated with logFerritin and serum iron (Table 1). Lastly, all the SNPs contained in our decaplex [37] were associated with at least one parameter, with the exception of rs4820268, which is in strong LD with rs855791 (both in *TMPRSS6*).

The SNP rs1375515 (*CACNA2D3*) was found to be associated with haemoglobin and logFerritin.

Table 1. Descriptive and linear regression for SNPs and biochemical parameters. Only significant results were shown. Effect model analyses do not apply for those SNPs that do not present individuals homozygous for the rare allele

Reference	SNPs	Location		Parameter	Model	Effect Genotype	Coef (CI 95%)	p
		chr	gene					
§	rs16826756	2	—	logMCV	Rec	G/G	-0.02 (-0.04, 0.00)	0.033
§	rs2673289	2	—	Haemoglobin	Rec	T/T	-0.47 (-0.84, -0.09)	0.014
				logFerritin	Rec	T/T	-0.11 (-0.22, -0.01)	0.033
				Haematocrit	Rec	T/T	-1.15 (-2.14, -0.16)	0.024
*	rs13089763	3	CACNA2D3	logFerritin		C/G	0.19 (0.02, 0.36)	0.03
§	rs1375515	3	CACNA2D3	Haemoglobin	Rec	G/G	-0.49 (-0.98, -0.01)	0.046
				logFerritin	Rec	G/G	-0.16 (-0.29, -0.02)	0.022
*	rs1868498	3	CACNA2D3	logFerritin	Dom	A/G-G/G	0.12 (0.01, 0.22)	0.027
				Serum Iron	Rec	G/G	60.67 (21.67, 99.67)	0.0025
*	rs1868505	3	CACNA2D3	logFerritin	Dom	A/G-A/A	-0.10 (-0.19, 0.00)	0.049
*	rs4974366	3	CACNA2D3	logFerritin	Dom	C/T-T/T	0.11 (0.01, 0.21)	0.038
				Serum Iron	Rec	T/T	76.61 (31.95, 121.26)	0.0009
*	rs7653648	3	CACNA2D3	logFerritin	Dom	C/G-G/G	0.09 (0.01, 0.17)	0.03
§	rs1799852	3	TF	Haemoglobin		C/T	-0.39 (-0.71, -0.07)	0.019
				Haematocrit		C/T	-1.09 (-1.96, -0.23)	0.014
				Transferrin		C/T	-25.45 (-39.29, -11.61)	0.0004
§	rs2280673	3	TF	Transferrin	Rec	A/A	19.96 (1.49, 38.42)	0.035
§	rs3811647	3	TF	Transferrin	Cod	G/A	21.29 (8.74, 33.84)	<0.0001
						A/A	51.48 (30.88, 72.07)	
[17,38]	rs11970772	6	CCND3	Serum Iron	Rec	A/A	-24.85 (-48.80, -0.89)	0.043
[38,41]	rs4895441	6	HBS1L	logFerritin	Rec	G/G	0.36 (0.11, 0.62)	0.0057
				Serum Iron	Dom	A/G-G/G	11.04 (2.14, 19.94)	0.016
				Transferrin	Rec	G/G	-44.35 (-84.67, -4.03)	0.032
				logMCV	Rec	G/G	0.03 (0.01, 0.06)	0.0053
§	rs1799945	6	HFE	Transferrin	Dom	G/C-G/G	-16.79 (-29.64, -3.95)	0.011
§	rs1800562	6	HFE	logFerritin		G/A	0.17 (0.00, 0.34)	0.048
				Transferrin		G/A	-39.35 (-64.67, -14.02)	0.0025
[38,39]	rs198846	6	HIST1H1T	Transferrin	Dom	A/G-A/A	-16.06 (-28.65, -3.47)	0.013
[38,40]	rs13194491	6	HIST1H2BJ	Haemoglobin		C/T	-0.60 (-1.18, -0.03)	0.04
				Haematocrit		C/T	-1.64 (-3.16, -0.11)	0.036
				Serum Iron		C/T	-19.49 (-36.35, -2.63)	0.024
[38,41,43,44]	rs17342717	6	SLC17A1	Transferrin	Dom	C/T	-21.84 (-38.63, -5.05)	0.011
[38,40]	rs12216125	6	TRIM38	Transferrin	Dom	C/T-T/T	-25.30 (-37.35, -13.24)	<0.0001
				logMCV	Rec	T/T	0.02 (0.00, 0.03)	0.0094
[38,42]	rs3184504	12	SH2B3	Haematocrit	Dom	C/T-T/T	0.92 (0.07, 1.76)	0.034
[38,39]	rs11089823	22	TMPRSS6	Haemoglobin	Dom	C/T-C/C	0.34 (0.04, 0.64)	0.025
				Haematocrit	Dom	C/T-C/C	0.95 (0.16, 1.74)	0.018
§	rs855791	22	TMPRSS6	Transferrin	Rec	T/T	15.43 (0.07, 30.78)	0.049

[] Literature

* Fine mapping

§ Decaplex

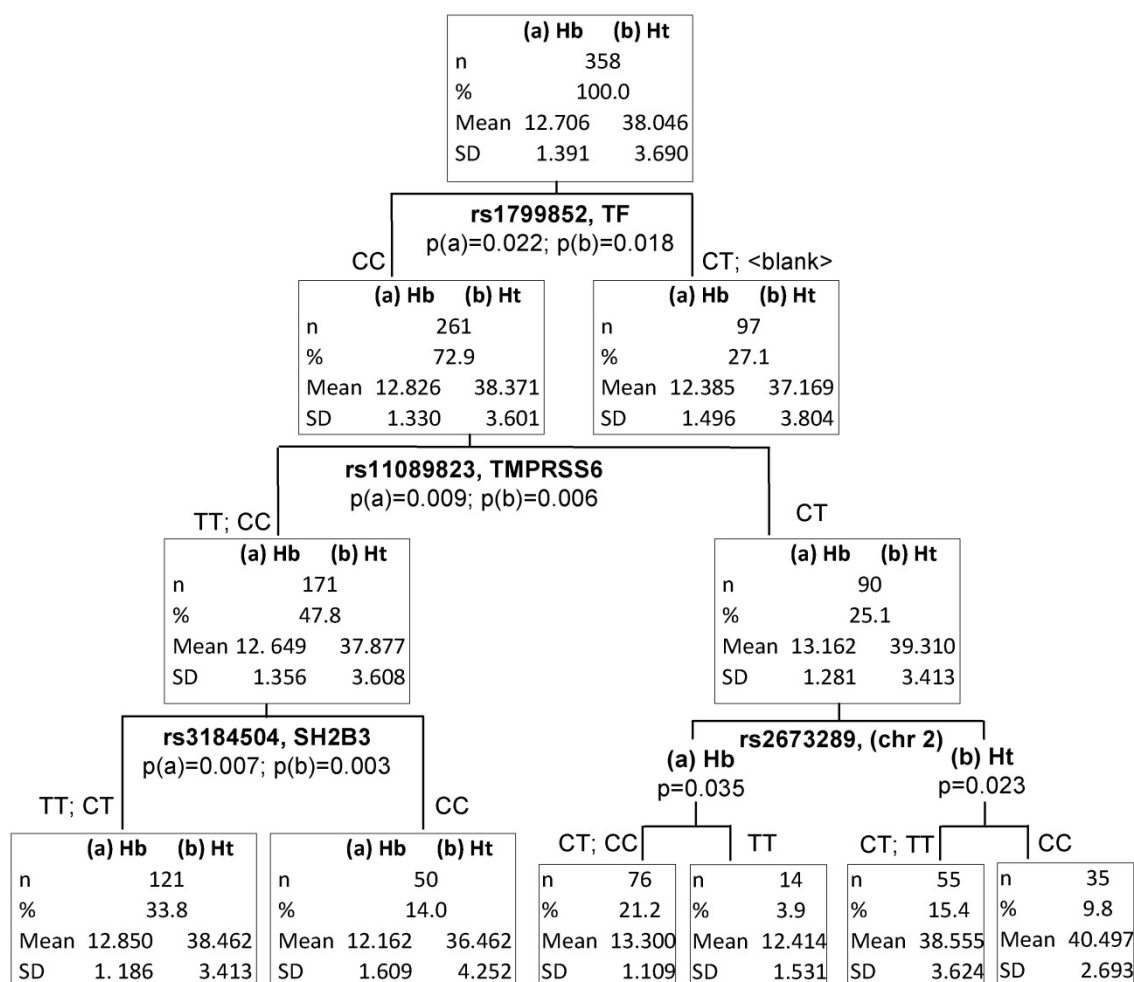


Fig. 1. Decision tree built for: (a) Haemoglobin (g/dl) and (b) Haematocrit (%), as dependant variable and the 45SNPs as independent variables or factors. Gene location of SNPs showed in capital letters. P-value: Significant value of the Chi-Square test that measures how well a SNP splits a sample according to the genotypes, %: Percentage of the total sample included in each node, blank: Missing values. CC individuals of rs11089823 were grouped with TT bearers; a statistical artefact due to small number of CC (six individuals)

All the SNPs that were found in the classification trees were also found in the regression analysis. As the decision trees for haemoglobin and haematocrit were built in a very similar way, both results are shown in Figure 1. The combination of CC of rs1799852 (*TF*), CT of rs11089823 (*TMPRSS6*) and CC/CT (in haemoglobin), CC (in haematocrit) of rs2673289 (chr 2) described the node with the highest level of both parameters. Conversely, the aggregate of CC of rs1799852 (*TF*), TT/CT of rs11089823 (*TMPRSS6*) and CC of rs3184504 (*SH2B3*) described

the group of samples (node) with the lowest level. The tree for logFerritin is depicted in Figure 2, where it can be seen that the SNP rs1375515 (*CACNA2D3*) was the one that most significantly split the variable into two groups. The entire sample could be described as follows: those individuals carrying AA/GA for rs1375515 (*CACNA2D3*) and GA for rs1800562 (*HFE*) presented the highest mean of logFerritin, whereas those carrying GG of rs1375515 (*CACNA2D3*) showed the

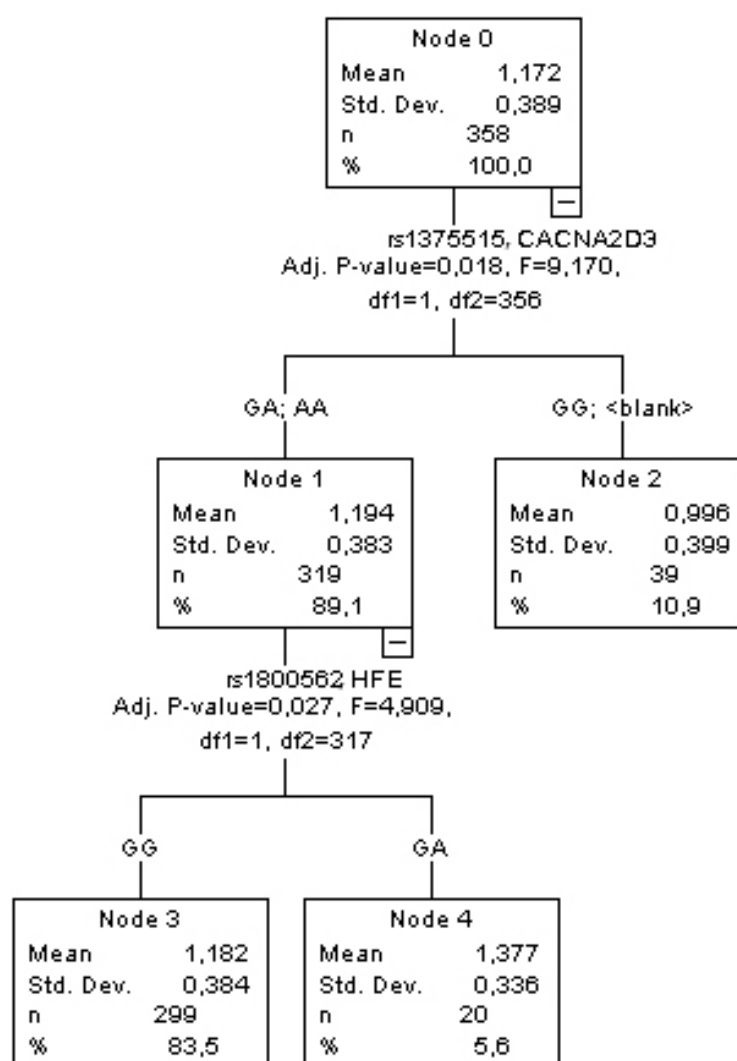


Fig. 2. Decision tree built for logFerritin (ng/ml) as dependant variable and the 45 SNPs as independent variables or factors. Gene location of SNPs showed in capital letters. P-value: Significant value of the Chi-Square test that measures how well a SNP splits a sample according to the genotypes, %: Percentage of the total sample included in each node, blank: Missing values

lowest values of this biochemical parameter. Figure 3 represents the tree for transferrin, illustrating how the sample was first split by the rs3811647 (*TF*) in what seems to be a co-dominant effect with a very high level of significance. The group carrying GG of this SNP yielded the lowest value, while the AA of rs3811647 and TT of 11089823 (*TMPRSS6*) yielded the highest. The decision tree for logMCV (Figure 1S) shows a node representing GG of rs16826756 (chr 2) with a lower mean than the general sample mean (1.909fl. vs 1.932fl.).

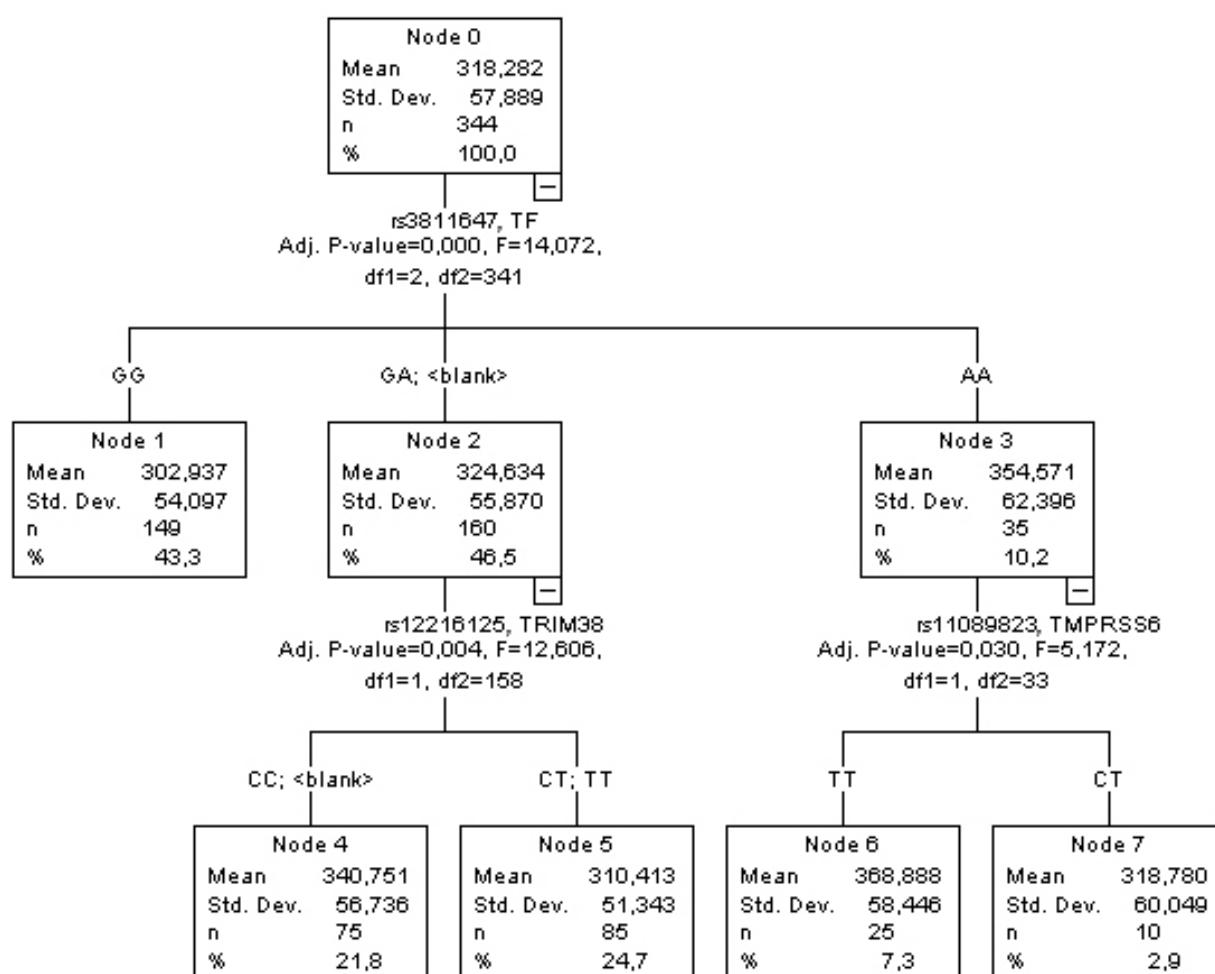


Fig. 3. Decision tree built for Transferrin as dependant variable and the 45 SNPs as independent variables or factors. Gene location of SNPs showed in capital letters. P-value: Significant value of the Chi-Square test that measures how well a SNP splits a sample according to the genotypes, %: Percentage of the total sample included in each node. This parameter was analysed in 344 individuals, blank: Missing values

Association of SNPs with iron-clinical phenotypes

The SNPs related to iron-clinical phenotypes are shown in Figure 4. A variant CT of rs1319441 (*HBS1L-MYB*) increased the probability of belonging to the anaemic group, in the same way as the variant GA of rs1800562 (*HFE*) did to the normal group. Therefore, the GA of rs1800562 almost doubled the proportion of the normal group in comparison with the general sample (36.0% vs 66.7%). On the opposite end of the distribution, the subsample with GG for rs1800562 (*HFE*) and CT for rs1319441 (*HBS1L-MYB*) more than doubled the proportion of the anaemic group compared with the general sample (58.3% vs 22.3%).

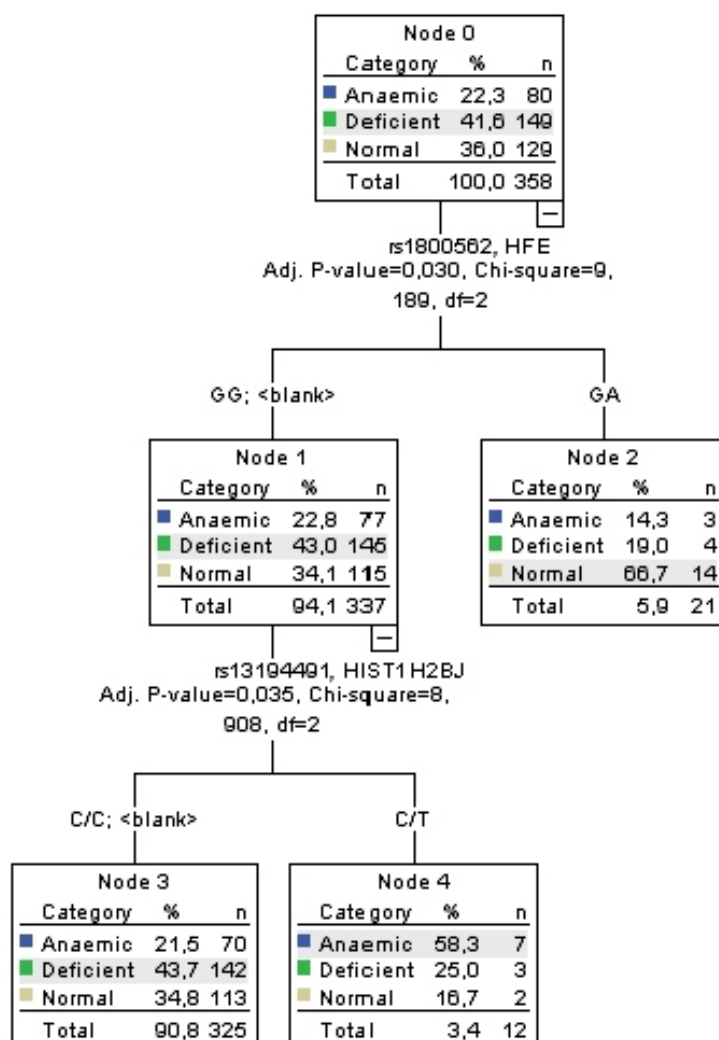


Fig. 4. Decision tree built for Iron-clinical phenotypes as dependant variable and the 45 SNPs as independent variables or factors. Gene location of SNPs showed in capital letters. P-value: Significant value of the Chi-Square test that measures how well a SNP splits a sample according to the genotypes, %: Percentage of the total sample included in each node, blank: Missing values

Logistic regression revealed the SNP variant that modified the odds of being anaemic (Table 2). The probability of being anaemic vs deficient/normal or of being anaemic/deficient vs normal increased for AG carriers of rs1868505 (*CACNA2D3*) by 1.690 times and by 3.067 times in CT carriers of rs13194491 (*HIST1H2B*). The GG genotype of rs48955441 (*HBS1L-MYB*) vs AA or AG had the highest protective effect (0.067 and 0.070, respectively). GA in rs1800562 (*HFE*) also decreased the risk of developing anaemia (0.169). This model with four SNPs correctly classified 46% of the individuals according to their phenotypes. Figure 5 presents a graphical illustration of the predictions made by the model. Overall, the figure explains the increased or decreased risk of being included in a given iron-clinical group in relation to the combined genotypes. For instance, the sixth genotype (**AA-GG-AG-CT**) was the one that predicted the highest probability of being anaemic and the lowest of being normal. In this example, as we have seen from other results, the less frequent variants of rs4895441 (*HBS1L-MYB*) (heterozygous and homozygous) and rs1800562 (*HFE*) exerted a protective effect in relation to iron status, whereas the rare variants of rs1314491 (*HIST1H2B*) and rs1868505 (*CACNA2D3*) had a harmful effect. Therefore, when the normal variants of the first two SNPs are combined with rare variants of the last two SNPs, the probability of belonging to the anaemic group is the highest. This rationale is valid for the rest of the genotypes.

Table 2. Back-step logistic regression model assuming iron-clinical phenotypes as the dependant variable and the SNPs as the factors

		df	Coef. Wald	Error	Chi- square	p	Comparison	Odds Ratio	CI 95%	
Intercept	1	1	-2.3127	0.4481	26.6380	<0.0001				
Intercept	2	1	-0.3949	0.4375	0.8148	0.3667				
rs4895441	GG	1	-1.7850	0.7190	6.1639	0.0130	GG vs AG	0.070	0.008	0.591
	AG	1	0.8724	0.3793	5.2905	0.0214	GG vs AA	0.067	0.008	0.560
							AG vs AA	0.961	0.624	1.480
rs1800562	GA	1	-0.8884	0.2712	10.7353	0.0011	GA vs GG	0.169	0.058	0.490
rs1868505	AG	1	0.2624	0.1228	45.690	0.0326	AG vs GG	1.690	1.045	2.735
rs13194491	CT	1	-0.5603	0.2409	54.117	0.0200	CT vs CC	3.067	1.193	7.884

Odds ratios inform of the probability of disease: being anaemic versus deficient/normal and anaemic/deficient versus normal. SNP rs1868505 is assumed to have a dominant effect, thus, the five AA individuals are included in the GA group. Location of the SNPs: *HBS1L-MYB*, *HFE*, *CACNA2D3*, *HBS1L-MYB* respectively

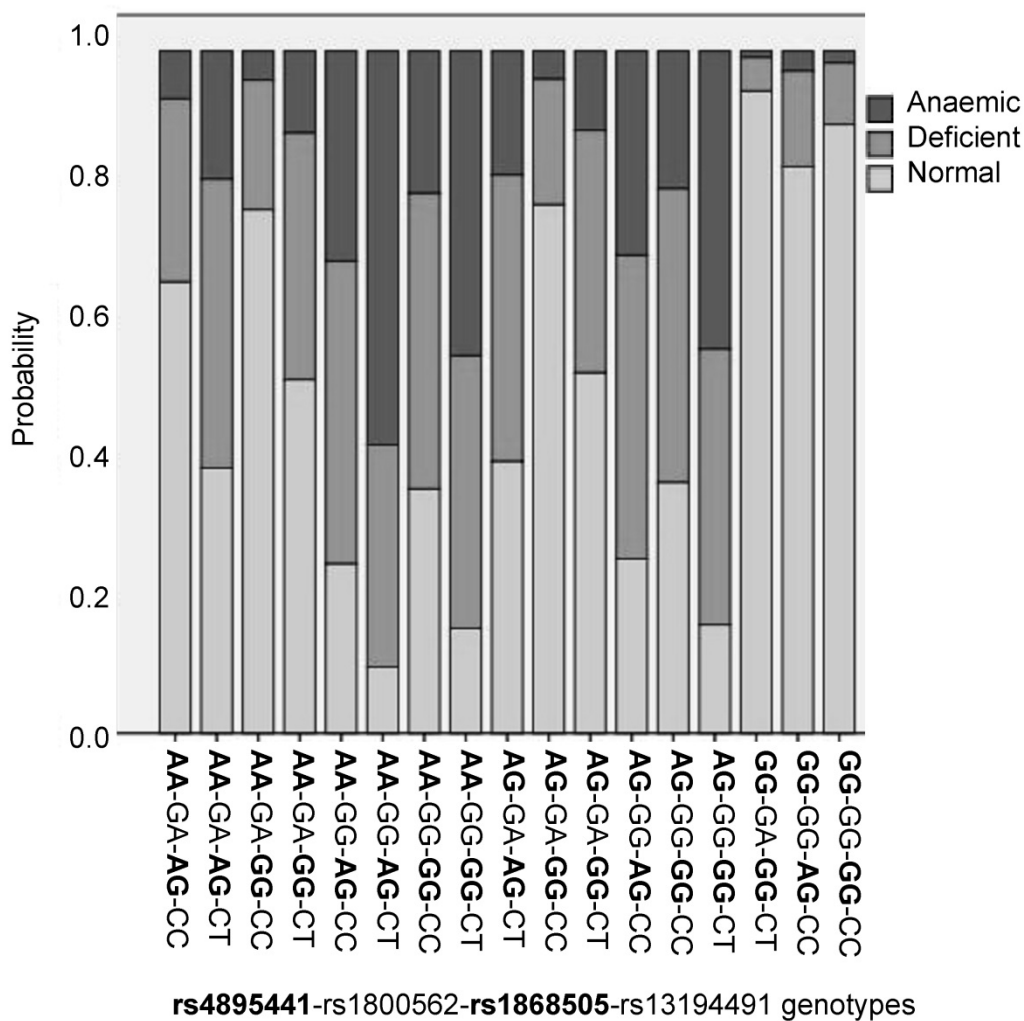


Fig. 5. Probabilities of presenting certain iron-clinical phenotypes regarding the combined genotypes of those SNPs significantly included in the back-step logistic regression. SNP rs1868505 is assumed to have a dominant effect, thus, the five AA individuals are included in the GA group. Location of the SNPs: *HBS1L-MYB*, *HFE*, *CACNA2D3*, *HBS1L-MYB*, respectively

4. Discussion

The present study has provided new insights into the genetic regulation of iron metabolism by testing the association of 45 SNPs with biochemical and iron-related parameters. Besides replicating the associations in our sample, we have tested some of them for the first time with iron-clinical phenotypes, obtaining significant results. Fine mapping of the genetic region of *CACNA2D3*, where

rs1375515 is located, has revealed a possible direct route of entry for iron through voltage dependant calcium genes, according to the significant association observed with ferritin, serum iron and iron-clinical phenotypes. Moreover, four SNPs explain a significant portion of the genetic variance in iron status, two of which are protective factors (rs1800562 *HFE*, rs4895441 *HBS1L-MYB*) while the other two are risk factors (rs1314491 *HIST1H2BJ*, rs1868505 *CACNA2D3*). .

Serum transferrin levels correlate inversely with the other parameters, as they are modulated according to the organism's iron needs [45-46]. Thus, SNPs generally present opposite effects for transferrin versus haemoglobin, ferritin, serum iron or MCV. The exception to this is rs1799852, as it reduces the levels not only of transferrin but also of haemoglobin and haematocrit, presumably by hampering normal transferrin activity or expression. This latter SNP, recently associated with iron status, together with rs3811647 and rs2280673, is situated in the *TF* gene (chr 3), influencing the levels of this protein [15-47]. The present study replicates all of these associations as previously reported elsewhere [18,48]. The rs12216125 (chr 6) shows a similar behaviour, negatively affecting transferrin and MCV. Ganesch et al. found that it was associated with MCV and haemoglobin [38], while Benyamin et al. [15] associated it with transferrin saturation. The present results suggest a more direct association with transferrin than with MCV, although the mechanism of this relationship remains unclear since this SNP is located in an intergenic region in 6p22.2.

As observed elsewhere, chromosome 6 seems to be very important as regards iron parameters [38,49], since the most significant SNPs found in this study were located on this chromosome. SNP rs4895441 is one of the intergenic variants of *HBS1L-MYB* 6q23 that are associated with elevated foetal haemoglobin [38,41,50]. *MYB* has been associated with the proliferation, survival and differentiation of haematopoietic progenitor cells [38]. The rare variant of this SNP generates an increase in ferritin, serum iron and logMCV, which is consistent with previous findings [41]. This could explain why this SNP presents a protective effect regarding iron status, as reported elsewhere [51]. C282Y is a well-known missense mutation of *HFE* (rs1800562) 6p22.2, responsible for hereditary

haemochromatosis when homozygous [9]. Several GWAS report this variant as a key regulator in erythrocyte traits, especially haemoglobin and haematocrit [17,38,40,42,43,49,52]. As in our previous results [21], we show here that this SNP is a protective factor for iron status. Rare variants of SNP rs13194491 in 6p21.33 behave as risk factors for iron status [38,40]. Located in *HIST1H2BJ*, a gene that expresses H2B type 1-J histone, its physiological implications are still unknown.

Fine mapping of the region that encodes the subunit of a voltage dependant calcium channel *CACNA2D3* (3p21.1) revealed several SNPs associated specifically with ferritin and serum iron. This region is associated with levels of transferrin receptor 1 (TfR1) [35]. *CACNA2D3* presents a regulatory function, modulating calcium current density and activation/inactivation of the kinetics of L-type calcium channels [53]. As regards the association found between rs1375515 and MCV, in our previous study [21] we discussed the implication of calcium channels, giving the relationship between ionic exchange, erythrocyte volume and iron pathologies [54]. However, the present data suggest a more direct relation with iron itself (see highly significant association of rs4974366 with serum iron). Along the same lines, several studies have reported voltage dependent calcium channels as a possible direct route for iron uptake in cardiomyocytes and neuron cells [30,31,55,56]. The opposite and seemingly independent effects found for these SNPs and their LD pattern are consistent with the regulatory function of the *CACNA2D3* subunit and reinforce its potential role as iron pathway. For instance, rs1375515 and rs1868505 are in LD (Table 2S) and decrease the levels of haemoglobin and ferritin, whereas rs49743366 and rs1868498 increase the levels of serum iron and ferritin. Ferritin not only stores intracellular iron, but as demonstrated in neuronal cell cultures, it can act as an extracellular transporter and could be involved in the regulation of iron uptake independently from transferrin [30]. This key protein is ubiquitously expressed and is up-regulated by serum iron through an iron responsive element (IRE). In iron overload conditions, Fe^{2+} has been reported to compete with Ca^{2+} on its way through calcium channels to the cytoplasm, causing cardiomyopathies or neuronal damage; therefore, blocking these channels with specific drugs is common in iron overload treatments [57], and may produce anaemia [54]. Most previous studies proving these channels

as iron pathways have been carried out in thalassemic or iron overload conditions, cultured cells and mice models. Conversely, according to the present results, we suggest that L-type calcium channels regulated by *CACNA2D3* could be a direct way of iron uptake in non-pathological subjects, via a possible ferritin-involved pathway, since all the SNPs in *CACNA2D3* are exclusively associated with ferritin and serum iron.

Numerous GWAS and other association studies have associated genetic loci with haematological traits, but very few have associated them with iron status [20,49]. Here, we present a genetic model that estimates the probabilities of belonging to a given iron-clinical phenotype. The model correctly classified a significant proportion of individuals, giving the heritability range values of haemoglobin and ferritin: 0.34-0.59 for haemoglobin and 0.20-0.29 for ferritin [12,13,58,59]. Moreover, the division of iron status into two categories instead of three, following WHO criteria [4], could improve predictive power, as could a larger sample.

In conclusion, this study replicates the association of SNPs with iron-related traits, and provides new information about the possible mechanisms of these relationships. Our results suggest a protective effect of two rare variants: the A allele of rs1800562 in *HFE*, and the G allele of rs4895441 in *HBS1L-MYB*. In contrast, the A allele of rs1868505 in *CACNA2D3* and the T allele of rs13194491 in *HIST1H2BJ* increase the risk of being anaemic in reproductive age women. Finally, we propose L-type calcium channels as a candidate pathway for direct iron uptake in physiological conditions. This hypothesis should be further investigated in future expression, molecular, animal model or other studies.

Acknowledgments

This study was supported by projects GR35/10 and AGL2009-11437. We would like to thank M. Carmen Bravo, Verónica Alonso, Ana Pérez, Ione Wright, Cristina Fernández, and Gema Moreno.

Bibliography

- [1] M.W. Hentze, M.U. Muckenthaler, and N.C. Andrews, Balancing acts: molecular control of mammalian iron metabolism. *Cell* 117 (2004) 285-97.
- [2] L.L. Dunn, Y. Suryo Rahmanto, and D.R. Richardson, Iron uptake and metabolism in the new millennium. *Trends Cell Biol* 17 (2007) 93-100.
- [3] G.J. Anderson, D.M. Frazer, and G.D. McLaren, Iron absorption and metabolism. *Curr Opin Gastroenterol* 25 (2009) 129-35.
- [4] B. de Benoist, E. McLean, I. Egli, and M. Cogswell, (Eds.), Worldwide prevalence of anaemia 1993-2005: WHO Global Database on Anaemia., World Health Organization, Geneva, 2008.
- [5] A.C. Crespo, B. Silva, L. Marques, E. Marcelino, C. Maruta, S. Costa, A. Timoteo, A. Vilares, F.S. Couto, P. Faustino, A.P. Correia, A. Verdelho, G. Porto, M. Guerreiro, A. Herrero, C. Costa, A. de Mendonca, L. Costa, and M. Martins, Genetic and biochemical markers in patients with Alzheimer's disease support a concerted systemic iron homeostasis dysregulation. *Neurobiol Aging* 35 777-85.
- [6] L. Toxqui, and M.P. Vaquero, Chronic Iron Deficiency as an Emerging Risk Factor for Osteoporosis: A Hypothesis. *Nutrients* 7 (2015) 2324-2344.
- [7] N. Abbaspour, R. Hurrell, and R. Kelishadi, Review on iron and its importance for human health. *J Res Med Sci* 19 (2014) 164-74.
- [8] R. Blanco-Rojo, L. Toxqui, A.M. Lopez-Parra, C. Baeza-Richer, A.M. Perez-Granados, E. Arroyo-Pardo, and M.P. Vaquero, Influence of diet, menstruation and genetic factors on iron status: a cross-sectional study in Spanish women of childbearing age. *Int J Mol Sci* 15 (2014) 4077-87.
- [9] J.N. Feder, A. Gnirke, W. Thomas, Z. Tsuchihashi, D.A. Ruddy, A. Basava, F. Dormishian, R. Domingo, Jr., M.C. Ellis, A. Fullan, L.M. Hinton, N.L. Jones, B.E. Kimmel, G.S. Kronmal, P. Lauer, V.K. Lee, D.B. Loeb, F.A. Mapa, E. McClelland, N.C. Meyer, G.A. Mintier, N. Moeller, T. Moore, E. Morikang, C.E. Prass, L. Quintana, S.M. Starnes, R.C. Schatzman, K.J. Brunke, D.T. Drayna, N.J. Risch, B.R. Bacon, and R.K. Wolff, A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* 13 (1996) 399-408.
- [10] S.G. Gehrke, A. Pietrangelo, M. Kascak, A. Braner, M. Eisold, H. Kulaksiz, T. Herrmann, U. Hebling, K. Bents, R. Gugler, and W. Stremmel, HJV gene mutations in European patients with juvenile hemochromatosis. *Clin Genet* 67 (2005) 425-8.
- [11] K.E. Finberg, Iron-refractory iron deficiency anemia. *Semin Hematol* 46 (2009) 378-86.
- [12] C. Garner, T. Tatu, J.E. Reittie, T. Littlewood, J. Darley, S. Cervino, M. Farrall, P. Kelly, T.D. Spector, and S.L. Thein, Genetic influences on F cells and other hematologic variables: a twin heritability study. *Blood* 95 (2000) 342-6.
- [13] J.C. Souto, A. Remacha, A. Buil, L. Almasy, J. Blangero, and J. Fontcuberta, Genetic determinants of iron metabolism plasma phenotypes and their relationship with risk of thrombosis. *Haematologica* 88 (2003) 1436-8.

- [14] O.T. Njajou, B.Z. Alizadeh, Y. Aulchenko, M.C. Zillikens, H.A. Pols, B.A. Oostra, D.W. Swinkels, and C.M. van Duijn, Heritability of serum iron, ferritin and transferrin saturation in a genetically isolated population, the Erasmus Rucphen Family (ERF) Study. *Hum Hered* 61 (2006) 222-8.
- [15] B. Benyamin, A.F. McRae, G. Zhu, S. Gordon, A.K. Henders, A. Palotie, L. Peltonen, N.G. Martin, G.W. Montgomery, J.B. Whitfield, and P.M. Visscher, Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels. *Am J Hum Genet* 84 (2009) 60-5.
- [16] B. Benyamin, T. Esko, J.S. Ried, A. Radhakrishnan, S.H. Vermeulen, M. Traglia, M. Gogele, D. Anderson, L. Broer, C. Podmore, J. Luan, Z. Kutalik, S. Sanna, P. van der Meer, T. Tanaka, F. Wang, H.J. Westra, L. Franke, E. Mihailov, L. Milani, J. Haldin, J. Winkelmann, T. Meitinger, J. Thiery, A. Peters, M. Waldenberger, A. Rendon, J. Jolley, J. Sambrook, L.A. Kiemeny, F.C. Sweep, C.F. Sala, C. Schwienbacher, I. Pichler, J. Hui, A. Demirkan, A. Isaacs, N. Amin, M. Steri, G. Waeber, N. Verweij, J.E. Powell, D.R. Nyholt, A.C. Heath, P.A. Madden, P.M. Visscher, M.J. Wright, G.W. Montgomery, N.G. Martin, D. Hernandez, S. Bandinelli, P. van der Harst, M. Uda, P. Vollenweider, R.A. Scott, C. Langenberg, N.J. Wareham, C. van Duijn, J. Beilby, P.P. Pramstaller, A.A. Hicks, W.H. Ouwehand, K. Oexle, C. Gieger, A. Metspalu, C. Camaschella, D. Toniolo, D.W. Swinkels, and J.B. Whitfield, Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis. *Nat Commun* 5 (2014) 4926.
- [17] N. Soranzo, T.D. Spector, M. Mangino, B. Kuhnle, A. Rendon, A. Teumer, C. Willenborg, B. Wright, L. Chen, M. Li, P. Salo, B.F. Voight, P. Burns, R.A. Laskowski, Y. Xue, S. Menzel, D. Altshuler, J.R. Bradley, S. Bumpstead, M.S. Burnett, J. Devaney, A. Doring, R. Elosua, S.E. Epstein, W. Erber, M. Falchi, S.F. Garner, M.J. Ghorri, A.H. Goodall, R. Gwilliam, H.H. Hakonarson, A.S. Hall, N. Hammond, C. Hengstenberg, T. Illig, I.R. Konig, C.W. Knouff, R. McPherson, O. Melander, V. Mooser, M. Nauck, M.S. Nieminen, C.J. O'Donnell, L. Peltonen, S.C. Potter, H. Prokisch, D.J. Rader, C.M. Rice, R. Roberts, V. Salomaa, J. Sambrook, S. Schreiber, H. Schunkert, S.M. Schwartz, J. Serbanovic-Canic, J. Sinisalo, D.S. Siscovick, K. Stark, I. Surakka, J. Stephens, J.R. Thompson, U. Volker, H. Volzke, N.A. Watkins, G.A. Wells, H.E. Wichmann, D.A. Van Heel, C. Tyler-Smith, S.L. Thein, S. Kathiresan, M. Perola, M.P. Reilly, A.F. Stewart, J. Erdmann, N.J. Samani, C. Meisinger, A. Greinacher, P. Deloukas, W.H. Ouwehand, and C. Gieger, A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat Genet* 41 (2009) 1182-90.
- [18] R. Blanco-Rojo, C. Baeza-Richer, A.M. Lopez-Parra, A.M. Perez-Granados, A. Brichs, S. Bertoncini, A. Buil, E. Arroyo-Pardo, J.M. Soria, and M.P. Vaquero, Four variants in transferrin and HFE genes as potential markers of iron deficiency anaemia risk: an association study in menstruating women. *Nutr Metab (Lond)* 8 (2011) 69.
- [19] C.E. McLaren, M.J. Emond, V.N. Subramaniam, P.D. Phatak, J.C. Barton, P.C. Adams, J.B. Goh, C.J. McDonald, L.W. Powell, L.C. Gurrin, K.J. Allen, D.A. Nickerson, T. Louie, G.A. Ramm, G.J. Anderson, and G.D. McLaren, Exome sequencing in HFE C282Y homozygous men with extreme phenotypes identifies a GNPAT variant associated with severe iron overload. *Hepatology* (2015).
- [20] P. An, Q. Wu, H. Wang, Y. Guan, M. Mu, Y. Liao, D. Zhou, P. Song, C. Wang, L. Meng, Q. Man, L. Li, J. Zhang, and F. Wang, TMPRSS6, but not TF, TFR2 or BMP2 variants are

- associated with increased risk of iron-deficiency anemia. *Hum Mol Genet* 21 (2012) 2124-31.
- [21] C. Baeza-Richer, R. Blanco-Rojo, A.M. Lopez-Parra, A. Brichs, S. Bertoncini, A.M. Perez-Granados, A. Buil, J.M. Soria, E. Arroyo-Pardo, and M.P. Vaquero, Identification of a novel quantitative trait nucleotide related to iron status in a calcium channel gene. *Dis Markers* 34 (2013) 121-9.
- [22] E. Beutler, C. Van Geet, D.M. te Loo, T. Gelbart, K. Crain, J. Truksa, and P.L. Lee, Polymorphisms and mutations of human TMPRSS6 in iron deficiency anemia. *Blood Cells Mol Dis* 44 (2009) 16-21.
- [23] D.J. Balding, A tutorial on statistical methods for population association studies. *Nat Rev Genet* 7 (2006) 781-91.
- [24] A. Salas, and A. Carracedo, [Studies of association in complex diseases: statistical problems related to the analysis of genetic polymorphisms]. *Rev Clin Esp* 207 (2007) 563-5.
- [25] S.J. Chanock, T. Manolio, M. Boehnke, E. Boerwinkle, D.J. Hunter, G. Thomas, J.N. Hirschhorn, G. Abecasis, D. Altshuler, J.E. Bailey-Wilson, L.D. Brooks, L.R. Cardon, M. Daly, P. Donnelly, J.F. Fraumeni, Jr., N.B. Freimer, D.S. Gerhard, C. Gunter, A.E. Guttmacher, M.S. Guyer, E.L. Harris, J. Hoh, R. Hoover, C.A. Kong, K.R. Merikangas, C.C. Morton, L.J. Palmer, E.G. Phimister, J.P. Rice, J. Roberts, C. Rotimi, M.A. Tucker, K.J. Vogan, S. Wacholder, E.M. Wijsman, D.M. Winn, and F.S. Collins, Replicating genotype-phenotype associations. *Nature* 447 (2007) 655-60.
- [26] B. Castellani, and J. Castellani, Data mining: qualitative analysis with health informatics data. *Qual Health Res* 13 (2003) 1005-18.
- [27] I. Dunham, An integrated encyclopedia of DNA elements in the human genome. *Nature* 489 (2012) 57-74.
- [28] M.T. Maurano, R. Humbert, E. Rynes, R.E. Thurman, E. Haugen, H. Wang, A.P. Reynolds, R. Sandstrom, H. Qu, J. Brody, A. Shafer, F. Neri, K. Lee, T. Kutayavin, S. Stehling-Sun, A.K. Johnson, T.K. Canfield, E. Giste, M. Diegel, D. Bates, R.S. Hansen, S. Neph, P.J. Sabo, S. Heimfeld, A. Raubitschek, S. Ziegler, C. Cotsapas, N. Sotoodehnia, I. Glass, S.R. Sunyaev, R. Kaul, and J.A. Stamatoyannopoulos, Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337 (2012) 1190-5.
- [29] F. Lang, Mechanisms and significance of cell volume regulation. *J Am Coll Nutr* 26 (2007) 613S-623S.
- [30] J.A. Gaasch, W.J. Geldenhuys, P.R. Lockman, D.D. Allen, and C.J. Van der Schyf, Voltage-gated calcium channels provide an alternate route for iron uptake in neuronal cell cultures. *Neurochem Res* 32 (2007) 1686-93.
- [31] N. Chattipakorn, S. Kumfu, S. Fucharoen, and S. Chattipakorn, Calcium channels and iron uptake into the heart. *World J Cardiol* 3 (2011) 215-8.
- [32] S. Bertoncini, R. Blanco-Rojo, C. Baeza, E. Arroyo-Pardo, M.P. Vaquero, and A.M. Lopez-Parra, A novel SNaPshot assay to detect genetic mutations related to iron metabolism. *Genet Test Mol Biomarkers* 15 (2013) 173-9.

- [33] P.I. de Bakker, R. Yelensky, I. Pe'er, S.B. Gabriel, M.J. Daly, and D. Altshuler, Efficiency and power in genetic association studies. *Nat Genet* 37 (2005) 1217-23.
- [34] J.C. Barrett, B. Fry, J. Maller, and M.J. Daly, Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21 (2005) 263-5.
- [35] A.F. Remacha, J.C. Souto, J.M. Soria, A. Buil, M.P. Sarda, M. Lathrop, J. Blangero, L. Almasy, and J. Fontcuberta, Genomewide linkage analysis of soluble transferrin receptor plasma levels. *Ann Hematol* 85 (2006) 25-8.
- [36] D. Collet, (Ed.), *Modelling Binary Data*, Chapman & Hall, London, 2003.
- [37] S. Bertoncini, R. Blanco-Rojo, C. Baeza, E. Arroyo-Pardo, M.P. Vaquero, and A.M. Lopez-Parra, A novel SNaPshot assay to detect genetic mutations related to iron metabolism. *Genet Test Mol Biomarkers* 15 (2011) 173-9.
- [38] S.K. Ganesh, N.A. Zakai, F.J. van Rooij, N. Soranzo, A.V. Smith, M.A. Nalls, M.H. Chen, A. Kottgen, N.L. Glazer, A. Dehghan, B. Kuhnel, T. Aspelund, Q. Yang, T. Tanaka, A. Jaffe, J.C. Bis, G.C. Verwoert, A. Teumer, C.S. Fox, J.M. Guralnik, G.B. Ehret, K. Rice, J.F. Felix, A. Rendon, G. Eiriksdottir, D. Levy, K.V. Patel, E. Boerwinkle, J.I. Rotter, A. Hofman, J.G. Sambrook, D.G. Hernandez, G. Zheng, S. Bandinelli, A.B. Singleton, J. Coresh, T. Lumley, A.G. Uitterlinden, J.M. Vangils, L.J. Launer, L.A. Cupples, B.A. Oostra, J.J. Zwaginga, W.H. Ouwehand, S.L. Thein, C. Meisinger, P. Deloukas, M. Nauck, T.D. Spector, C. Gieger, V. Gudnason, C.M. van Duijn, B.M. Psaty, L. Ferrucci, A. Chakravarti, A. Greinacher, C.J. O'Donnell, J.C. Witteman, S. Furth, M. Cushman, T.B. Harris, and J.P. Lin, Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nat Genet* 41 (2009) 1191-8.
- [39] J.C. Chambers, W. Zhang, Y. Li, J. Sehmi, M.N. Wass, D. Zabaneh, C. Hoggart, H. Bayele, M.I. McCarthy, L. Peltonen, N.B. Freimer, S.K. Srail, P.H. Maxwell, M.J. Sternberg, A. Ruukonen, G. Abecasis, M.R. Jarvelin, J. Scott, P. Elliott, and J.S. Kooner, Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. *Nat Genet* 41 (2009) 1170-2.
- [40] B. Benyamin, M.A. Ferreira, G. Willemsen, S. Gordon, R.P. Middelberg, B.P. McEvoy, J.J. Hottenga, A.K. Henders, M.J. Campbell, L. Wallace, I.H. Frazer, A.C. Heath, E.J. de Geus, D.R. Nyholt, P.M. Visscher, B.W. Penninx, D.I. Boomsma, N.G. Martin, G.W. Montgomery, and J.B. Whitfield, Common variants in TMPRSS6 are associated with iron status and erythrocyte volume. *Nat Genet* 41 (2009) 1173-5.
- [41] I.J. Kullo, K. Ding, H. Jouni, C.Y. Smith, and C.G. Chute, A genome-wide association study of red blood cell traits using the electronic medical record. *PLoS One* 5 (2010).
- [42] K.S. Lo, J.G. Wilson, L.A. Lange, A.R. Folsom, G. Galarneau, S.K. Ganesh, S.F. Grant, B.J. Keating, S.A. McCarroll, E.R. Mohler, 3rd, C.J. O'Donnell, W. Palmas, W. Tang, R.P. Tracy, A.P. Reiner, and G. Lettre, Genetic association analysis highlights new loci that modulate hematological trait variation in Caucasians and African Americans. *Hum Genet* 129 (2010) 307-17.
- [43] K. Oexle, J.S. Ried, A.A. Hicks, T. Tanaka, C. Hayward, M. Bruegel, M. Gogele, P. Lichtner, B. Muller-Myhsok, A. Doring, T. Illig, C. Schwienbacher, C. Minelli, I. Pichler, G.M. Fiedler, J. Thiery, I. Rudan, A.F. Wright, H. Campbell, L. Ferrucci, S. Bandinelli, P.P. Pramstaller, H.E. Wichmann, C. Gieger, J. Winkelmann, and T. Meitinger, Novel

- association to the proprotein convertase PCSK7 gene locus revealed by analysing soluble transferrin receptor (sTfR) levels. *Hum Mol Genet* 20 (2011) 1042-7.
- [44] I. Pichler, C. Minelli, S. Sanna, T. Tanaka, C. Schwienbacher, S. Naitza, E. Porcu, C. Pattaro, F. Busonero, A. Zanon, A. Maschio, S.A. Melville, M. Grazia Piras, D.L. Longo, J. Guralnik, D. Hernandez, S. Bandinelli, E. Aigner, A.T. Murphy, V. Wroblewski, F. Marroni, I. Theurl, C. Gnewuch, E. Schadt, M. Mitterer, D. Schlessinger, L. Ferrucci, D.R. Witcher, A.A. Hicks, G. Weiss, M. Uda, and P.P. Pramstaller, Identification of a common variant in the TFR2 gene implicated in the physiological regulation of serum iron levels. *Hum Mol Genet* 20 (2011) 1232-40.
- [45] A. Rolfs, I. Kvietikova, M. Gassmann, and R.H. Wenger, Oxygen-regulated transferrin expression is mediated by hypoxia-inducible factor-1. *J Biol Chem* 272 (1997) 20055-62.
- [46] M.F. Macedo, E. Cruz, R. Lacerda, G. Porto, and M. de Sousa, Low serum transferrin levels in HFE C282Y homozygous subjects are associated with low CD8(+) T lymphocyte numbers. *Blood Cells Mol Dis* 35 (2005) 319-25.
- [47] W.N. Gichohi-Wainaina, A. Melse-Boonstra, D.W. Swinkels, M.B. Zimmermann, E.J. Feskens, and G.W. Towers, Common Variants and Haplotypes in the TF, TNF-alpha, and TMPRSS6 Genes Are Associated with Iron Status in a Female Black South African Population. *J Nutr* (2015).
- [48] R. Blanco-Rojo, H.K. Bayele, S.K. Srail, and M.P. Vaquero, Intronic SNP rs3811647 of the human transferrin gene modulates its expression in hepatoma cells. *Nutr Hosp* 27 (2012) 2142-5.
- [49] C.E. McLaren, C.P. Garner, C.C. Constantine, S. McLachlan, C.D. Vulpe, B.M. Snively, V.R. Gordeuk, D.A. Nickerson, J.D. Cook, C. Leiendecker-Foster, K.B. Beckman, J.H. Eckfeldt, L.F. Barcellos, J.A. Murray, P.C. Adams, R.T. Acton, A.A. Killeen, and G.D. McLaren, Genome-wide association study identifies genetic loci associated with iron deficiency. *PLoS One* 6 (2011) e17390.
- [50] R. Stadhouders, S. Aktuna, S. Thongjuea, A. Aghajanirefah, F. Pourfarzad, W. van Ijcken, B. Lenhard, H. Rooks, S. Best, S. Menzel, F. Grosveld, S.L. Thein, and E. Soler, HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. *J Clin Invest* 124 (2014) 1699-710.
- [51] V.G. Sankaran, and S.H. Orkin, The switch from fetal to adult hemoglobin. *Cold Spring Harb Perspect Med* 3 a011643.
- [52] T. Tanaka, C.N. Roy, W. Yao, A. Matteini, R.D. Semba, D. Arking, J.D. Walston, L.P. Fried, A. Singleton, J. Guralnik, G.R. Abecasis, S. Bandinelli, D.L. Longo, and L. Ferrucci, A genome-wide association analysis of serum iron concentrations. *Blood* 115 94-6.
- [53] C.A. Gurnett, M. De Waard, and K.P. Campbell, Dual function of the voltage-dependent Ca²⁺ channel alpha 2 delta subunit in current stimulation and subunit interaction. *Neuron* 16 (1996) 431-40.
- [54] V.L. Lew, Z. Etzion, and R.M. Bookchin, Dehydration response of sickle cells to sickling-induced Ca(++) permeabilization. *Blood* 99 (2002) 2578-85.

- [55] S. Kumfu, S. Chattipakorn, K. Chinda, S. Fucharoen, and N. Chattipakorn, T-type calcium channel blockade improves survival and cardiovascular function in thalassemic mice. *Eur J Haematol* 88 (2012) 535-48.
- [56] G.Y. Oudit, H. Sun, M.G. Trivieri, S.E. Koch, F. Dawood, C. Ackerley, M. Yazdanpanah, G.J. Wilson, A. Schwartz, P.P. Liu, and P.H. Backx, L-type Ca^{2+} channels provide a major pathway for iron entry into cardiomyocytes in iron-overload cardiomyopathy. *Nat Med* 9 (2003) 1187-94.
- [57] K. Sugishita, M. Asakawa, S. Usui, and T. Takahashi, A case of iron overload cardiomyopathy: beneficial effects of iron chelating agent and calcium channel blocker on left ventricular dysfunction. *Int Heart J* 50 (2009) 829-38.
- [58] J.B. Whitfield, and N.G. Martin, Genetic and environmental influences on the size and number of cells in the blood. *Genet Epidemiol* 2 (1985) 133-44.
- [59] C. Sala, M. Ciullo, C. Lanzara, T. Natile, S. Bione, R. Massacane, P. d'Adamo, P. Gasparini, D. Toniolo, and C. Camaschella, Variation of hemoglobin levels in normal Italian populations from genetic isolates. *Haematologica* 93 (2008) 1372-5.

Supplementary material

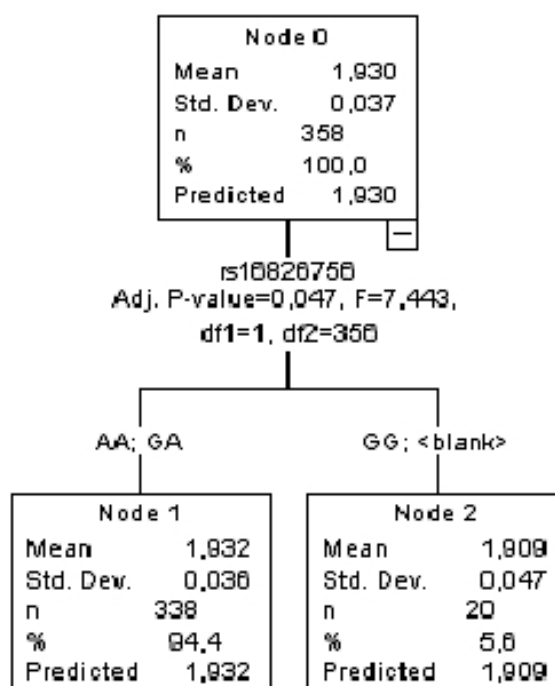


Fig 1S. Decission tree built for MCV(fl.) as dependent variable and the 45 SNPs as independent variables or factors. P-value: Significant value of the Chi-Square test that measures how well a SNP splits a sample according to the genotypes, %: Percentage of the total sample included in each node, blank: Missing values

Table 1S. Descriptive and linear regression for all the SNPs and biochemical parameters

Reference	SNPs	Location chr	gene	Parameter	Model	Effect Genotype	p
[49]	rs2698530	2	—				
[38,42]	rs10224002	7	<i>PRKAG2</i>				
[17,42,44]	rs7385804	7	<i>TFR2</i>				
[17,38]	rs11065987	12	—				
[38,52]	rs12435835	14	<i>MAX</i>				
[17,38]	rs11066301	12	<i>PTPN11</i>				
[38]	rs17630235	12	<i>TRAFD1</i>				
[38,42]	rs2782	22	<i>NCAPH2</i>				
[38,39]	rs11089823	22	<i>TMPRSS6</i>	Haemoglobin	Dom	C/T-C/C	0.025
				Haematocrit	Dom	C/T-C/C	0.018
[38,43]	rs2413450	22	<i>TMPRSS6</i>				
[49]	rs987710	22					
[38,40]	rs2274089	6	<i>LRRC16</i>				
[17,38]	rs11970772	6	<i>CCND3</i>	Serum Iron	Rec	A/A	0.043
[38,40]	rs12216125	6	<i>TRIM38</i>	Transferrin	Dom	C/T-T/T	<0.0001
				logMCV	Rec	T/T	0.0094
[38,40]	rs13194491	6	<i>HIST1H2BJ</i>	Haemoglobin		C/T	0.04
				Haematocrit		C/T	0.036
				Serum Iron		C/T	0.024
[38,41,43,44]	rs17342717	6	<i>SLC17A1</i>	Transferrin	Dom	C/T	0.011
[38,39]	rs198846	6	<i>HIST1H1T</i>	Transferrin	Dom	A/G-A/A	0.013
[38,42]	rs3184504	12	<i>SH2B3</i>	Haematocrit	Dom	C/T-T/T	0.034
[38,41]	rs4895441	6	<i>HBS1L</i>	logFerritin	Rec	G/G	0.0057
				Serum Iron	Dom	A/G-G/G	0.016
				Transferrin	Rec	G/G	0.032
				logMCV	Rec	G/G	0.0053
*	rs9841318	3	<i>CACNA2D3</i>				
*	rs9818236	3	<i>CACNA2D3</i>				
*	rs9311528	3	<i>CACNA2D3</i>				
*	rs7644514	3	<i>CACNA2D3</i>				
*	rs7638271	3	<i>CACNA2D3</i>				
*	rs6803199	3	<i>CACNA2D3</i>				
*	rs6790309	3	<i>CACNA2D3</i>				
*	rs4974369	3	<i>CACNA2D3</i>				
*	rs13088989	3	<i>CACNA2D3</i>				
*	rs1992964	3	<i>CACNA2D3</i>				
*	rs1449330	3	<i>CACNA2D3</i>				
*	rs13089763	3	<i>CACNA2D3</i>	logFerritin		C/G	0.03
*	rs1868498		<i>CACNA2D3</i>	logFerritin	Dom	A/G-G/G	0.027
				Serum Iron	Rec	G/G	0.0025
*	rs1868505	3	<i>CACNA2D3</i>	logFerritin	Dom	A/G-A/A	0.049
*	rs4974366		<i>CACNA2D3</i>	logFerritin	Dom	C/T-T/T	0.038
				Serum Iron	Rec	T/T	0.0009
*	rs7653648	3	<i>CACNA2D3</i>	logFerritin	Dom	C/G-G/G	0.03
[] Literature							
* Fine mapping							
§ Decaplex							

Table 1S cont. Descriptive and linear regression for all the SNPs and biochemical parameters

Reference	SNPs	Location chr	gene	Parameter	Model	Effect Genotype	p
§	rs855791	22	TMPRSS6	Transferrin	Rec	T/T	0.049
§	rs1799852	3	TF	Haemoglobin		C/T	0.019
				Haematocrit		C/T	0.014
				Transferrin		C/T	0.0004
§	rs2280673	3	TF	Transferrin	Rec	A/A	0.035
§	rs1800562	6	HFE	logFerritin		G/A	0.048
				Transferrin		G/A	0.0025
§	rs3811647	3	TF	Transferrin	Cod	G/A	<0.0001
						A/A	
§	rs2673289	2	—	Haemoglobin	Rec	T/T	0.014
				logFerritin	Rec	T/T	0.033
				Haematocrit	Rec	T/T	0.024
§	rs1375515	3	CACNA2D3	Haemoglobin	Rec	G/G	0.046
				logFerritin	Rec	G/G	0.022
§	rs1799945		HFE	Transferrin	Dom	G/C-G/G	0.011
§	rs16826756	2	—	logMCV	Rec	G/G	0.033
[] Literature							
* Fine mapping							
§ Decaplex							

Table 2S. Matrix of pairwise values of LD (r), among SNPs from fine mapping (rs13735515 included)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
rs13088989 (1)	.	0.4788	0.0375	0.0265	0.0023	0.05	0.0013	0.0363	0.0156	0.0924	0.0082	0.0762	0.1222	0.0199	0.1899	0.0178	0.057
rs13089763 (2)	.	.	0.1027	0.032	0.0659	0.1072	0.058	0.0623	0.037	0.1934	0.003	0.0881	0.2555	0.0203	0.1439	0.0458	0.0435
rs1449330 (3)	.	.	.	0.5444	0.2289	0.7981	0.5694	0.4726	0.1312	0.5496	0.0863	0.545	0.7592	0.6527	0.2601	0.642	0.2431
rs1868498 (4)	0.1297	0.5314	0.9696	0.1154	0.0742	0.3784	0.0145	0.3526	0.4753	0.0942	0.1154	0.1311	0.2113
rs1868505 (5)	0.2396	0.1299	0.1738	0.0588	0.4298	0.0448	0.4028	0.2605	0.1517	0.8559	0.1491	0.5411
rs1992964 (6)	0.5439	0.7252	0.1374	0.6965	0.0824	0.6664	0.8851	0.5587	0.2721	0.58	0.38
rs4974366 (7)	0.1159	0.0743	0.379	0.0145	0.3526	0.4758	0.0946	0.1478	0.1313	0.212
rs4974369 (8)	0.0994	0.5053	0.0278	0.4888	0.6487	0.7825	0.1972	0.7706	0.6113
rs6790309 (9)	0.1974	0.0067	0.204	0.1494	0.0867	0.0819	0.0853	0.1534
rs6803199 (10)	0.057	0.8646	0.7135	0.404	0.3709	0.4102	0.7144
rs7638271 (11)	0.055	0.0757	0.0178	0.0319	0.1328	0.0737
rs7644514 (12)	0.5859	0.3906	0.2976	0.4083	0.7173
rs7653648 (13)	0.5726	0.2041	0.5627	0.343
rs9311528 (14)	0.1724	0.9224	0.4888
rs9818236 (15)	0.1696	0.442
rs9841318 (16)	0.5257
rs1375515 (17)

**4.3. “Frequency distribution and detection of
mutations in genes implicated in iron deficiency
in women”**

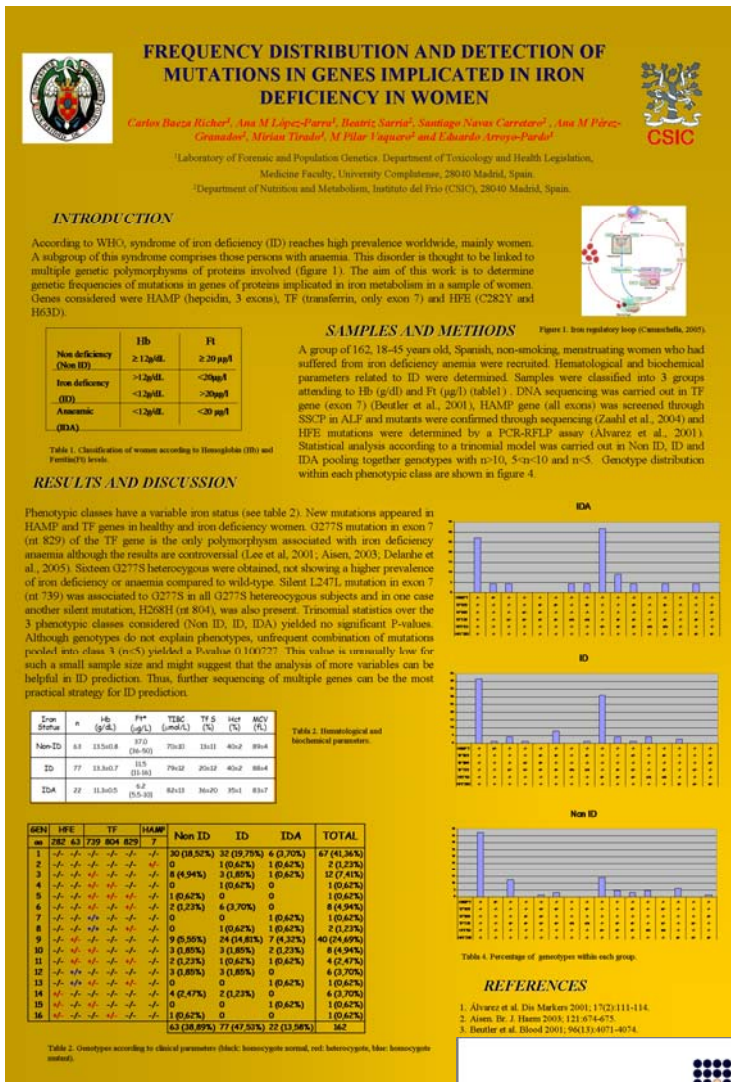
2006

Resultados complementarios

No publicado

Presentados en el 9º congreso *Portugaliæ Genetica*
- Genes and Populations in Health and Disease -
16-18 Marzo 2006, Porto (Portugal)

(Poster con defensa oral, Carlos Baeza Richer)



IPATIMUP
 Instituto de Patologia e Imunologia Molecular
 da Universidade do Porto

PORTUGALIAE GENETICA
 9th Edition

ATTGCGATAACCCGAATAC
 ATTCGGGATACCCGAATAC
 ATTCGCGATAACCCGAATAC
 ATTCGCGATAACCCGAATAC
 ATTCGCGATAACCCGAATAC

GENES AND POPULATIONS IN HEALTH AND DISEASE
 16 - 18 March 2006
 Porto, Portugal

4.3 “Frequency distribution and detection of mutations in genes implicated in iron deficiency in women” Presentación a congreso

Carlos Baeza Richer¹, Ana M López-Parra¹, Beatriz Sarria², Santiago Navas Carretero², Ana M Pérez-Granados², Mirian Tirado¹, M Pilar Vaquero² and Eduardo Arroyo-Pardo¹

¹Laboratory of Forensic and Population Genetics. Department of Toxicology and Health Legislation, Medicine Faculty, University Complutense, 28040 Madrid, Spain.

²Department of Nutrition and Metabolism, Instituto del Frío (CSIC), 28040 Madrid, Spain.

Author for correspondence: Carlos Baeza Richer, e-mail: cbaezaricher@med.ucm.es

Resumen

Introducción: De acuerdo con la organización mundial de la salud (OMS), el síndrome de la deficiencia de hierro presenta una alta prevalencia en la población mundial. La anemia ferropénica representa un subgrupo de este síndrome. Este desorden parece ser de herencia compleja, con lo que se espera una mayor prevalencia de mutaciones en la población afectada. **Objetivo:** Este trabajo tiene por objeto determinar las frecuencias de las mutaciones de proteínas implicadas en el metabolismo del hierro, en un grupo de mujeres españolas clasificadas por su estado de hierro. Los genes considerados fueron la hepcidina (*HAMP*, tres exones), la transferrina (*TF*, exón 7) y el gen *HFE* (C282Y y H63D). **Métodos:** Se reclutó un grupo de 162 mujeres, de entre 18 y 45 años, no fumadoras, y con antecedentes de anemia ferropénica. En estas muestras se analizaron parámetros bioquímicos y hematológicos relacionados con el metabolismo del hierro y se clasificaron las muestras en tres grupos de acuerdo a los valores de hemoglobina (Hb, g/dl) y ferritina (Ft, µg/l): sanas (Hb ≥12, Ft ≥20), deficientes (Hb≥12 and Ft<20 ó Hb<12 and Ft≥20) y anémicas (Hb <12 and Ft<20). Se secuenció el exón 7 de la transferrina, y se analizó la hepcidina por SSCP, confirmándose las posibles mutaciones mediante secuenciación. Las mutaciones del gen *HFE* se analizaron por PCR-RFLP. **Resultados:** Se detectaron nuevas mutaciones en el gen de la hepcidina y la transferrina en mujeres sanas y con deficiencia. Sólo se detectaron en homocigosis las mutaciones L247L (*TF*) y H63D (*HFE*). **Conclusión:** Las

regiones exónicas de *HAMP*, *TF* (exón 7) y *HFE* parecen muy conservadas en humanos aunque se pueden encontrar algunas mutaciones.

Abstract

Background: According to WHO, syndrome of iron deficiency reaches high prevalence worldwide, mainly women. A subgroup of this syndrome comprises those persons with anaemia. This disorder is thought to be linked to multiple genetic polymorphisms. Thus, individuals with iron deficiency are supposed to have a higher prevalence of mutations in genes involved in iron metabolism. The detection of some of these polymorphisms, their linkage and the distribution of frequencies is here addressed. **Aim:** In this work, we determine genetic frequencies of mutations in genes of proteins implicated in iron metabolism in a sample of women. Genes considered were *HAMP* (hepcidin, 3 exons), *TF* (transferrin, only exon 7) and *HFE* (C282Y and H63D). Description of new haplotypes was carried out. **Methods:** A group of 162, 18-45 years old, Spanish, non-smoking, menstruating women who had suffered from iron deficiency anemia or had a family history of anemia, were recruited. In these subjects, hematological and biochemical parameters related to iron metabolism were determined. Samples were classified into three groups attending to Hb (g/dl) and Ft ($\mu\text{g/l}$): healthy (Hb ≥ 12 , Ft ≥ 20), deficient (Hb ≥ 12 and Ft < 20 or Hb < 12 and Ft ≥ 20) and anaemic (Hb < 12 and Ft < 20). DNA sequencing was carried out in *TF* gene (exon 7). *HAMP* gene (all exons) was screened through SSCP and mutants were confirmed through sequencing. *HFE* mutations were determined by a PCR-RFLP assay. **Results:** New mutations appeared in *HAMP* and *TF* genes in healthy and iron deficiency women. Only mutations L247L (*TF*) and H63D (*HFE*) were detected in homozygosis. Genetic frequencies were reported for all the mutations studied in the 3 groups. **Conclusion:** Sequences of *HAMP*, *TF* (exon 7) and *HFE* seem to be highly conserved in humans, though some mutations can be observed.

Introduction

According to WHO, syndrome of iron deficiency (ID) reaches high prevalence worldwide, mainly women. A subgroup of this syndrome comprises those persons with anaemia. This disorder is thought to be linked to multiple genetic polymorphisms of proteins involved (Figure 1). The aim of this work is to determine genetic frequencies of mutations in genes of proteins implicated in iron metabolism in a sample of women. Genes considered were HAMP (hepcidin, 3 exons), TF (transferrin, only exon 7) and HFE (C282Y and H63D).

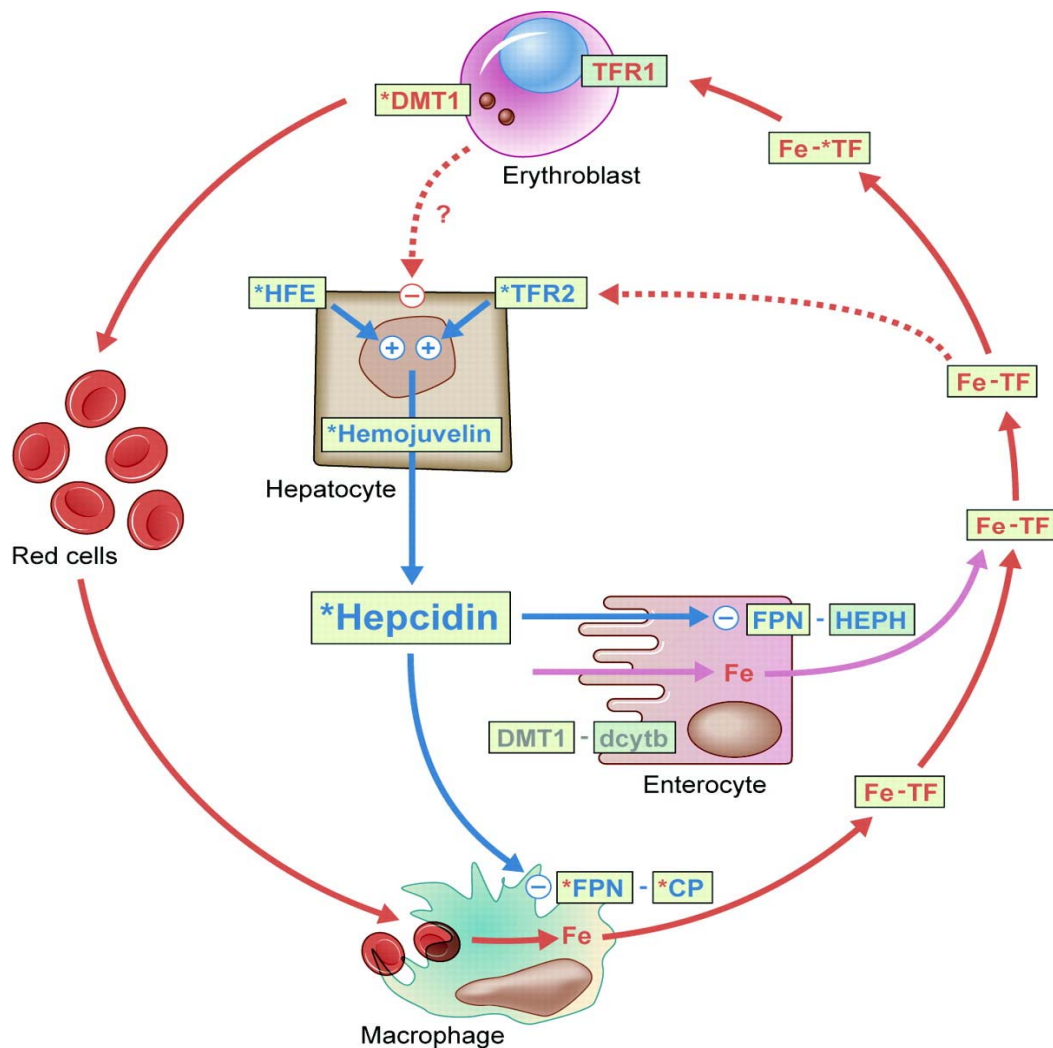


Figure 1. Iron regulatory loop (Camaschella, 2005)

Samples and Methods

A group of 162, 18-45 years old, Spanish, non-smoking, menstruating women who had suffered from iron deficiency anemia were recruited. Hematological and biochemical parameters related to ID were determined. Samples were classified into 3 groups attending to Hb (g/dl) and Ft ($\mu\text{g/l}$) (Table1). DNA sequencing was carried out in TF gene (exon 7) (Beutler et al., 2001), HAMP gene (all exons) was screened through SSCP in ALF and mutants were confirmed through sequencing (Zaahl et al., 2004) and HFE mutations were determined by a PCR-RFLP assay (Álvarez et al., 2001). Statistical analysis according to a trinomial model was carried out in Non ID, ID and IDA pooling together genotypes with $n > 10$, $5 < n < 10$ and $n < 5$. Genotype distribution within each phenotypic class is shown in Figure 2.

Table 1. Classification of women according to Hemoglobin (Hb) and Ferritin (Ft) levels

	Hb	Ft
Non deficiency (Non ID)	$\geq 12 \text{ g/dL}$	$\geq 20 \mu\text{g/L}$
Iron deficiency (ID)	$> 12 \text{ g/dL}$ $< 12 \text{ g/dL}$	$< 20 \mu\text{g/L}$ $> 20 \mu\text{g/L}$
Anaemic (IDA)	$< 12 \text{ g/dL}$	$< 20 \mu\text{g/L}$

Results and discussion

Phenotypic classes have a variable iron status (see Table 2). New mutations appeared in HAMP and TF genes in healthy and iron deficiency women. G277S mutation in exon 7 (nt 829) of the TF gene is the only polymorphism associated with iron deficiency anaemia although the results are controversial (Lee et al, 2001; Aisen, 2003; Delanhe et al., 2005). Sixteen G277S heterocygous were obtained, not showing a higher prevalence of iron deficiency or anaemia compared to wild-type. Silent L247L mutation in exon 7 (nt 739) was associated to G277S in all G277S heterocygous subjects and in one case another silent mutation, H268H (nt 804), was also present. Trinomial statistics over the 3 phenotypic classes considered (Non ID, ID, IDA) yielded no significant P-values. Although genotypes do not explain phenotypes (Table 3, Figure 2), unfrequent combination of mutations pooled into class 3 (n<5) yielded a P-value 0.100727. This value is unusually low for such a small sample size and might suggest that the analysis of more variables can be helpful in ID prediction. Thus, further sequencing of multiple genes can be the most practical strategy for ID prediction.

Table 2. Hematological and biochemical parameters

Iron Status	n	Hb (g/dL)	Ft (µg/L)	TIBC (µmol/L)	Tf S (%)	Hct (%)	MCV (fL)
Non-ID	63	13.5±0.8	37.0 (36-50)	70±10	13±11	40±2	89±4
ID	77	13.3±0.7	11.5 (11-16)	79±12	20±12	40±2	88±4
IDA	22	11.3±0.5	6.2 (5.5-10)	82±13	36±20	35±1	83±7

Table 3. Genotypes according to clinical parameters (black: homocygote normal, red: heterocygote, blue: homocygote mutant)

GEN	HFE		TF			HAMP	Non ID	ID	IDA	TOTAL
aa	282	63	739	804	829	7				
1	-/-	-/-	-/-	-/-	-/-	-/-	30 (18,52%)	32 (19,75%)	6 (3,70%)	67 (41,36%)
2	-/-	-/-	-/-	-/-	-/-	+/-	0	1 (0,62%)	1 (0,62%)	2 (1,23%)
3	-/-	-/-	+/-	-/-	-/-	-/-	8 (4,94%)	3 (1,85%)	1 (0,62%)	12 (7,41%)
4	-/-	-/-	+/-	+/-	-/-	-/-	0	1 (0,62%)	0	1 (0,62%)
5	-/-	-/-	+/-	+/-	+/-	-/-	1 (0,62%)	0	0	1 (0,62%)
6	-/-	-/-	+/-	-/-	+/-	-/-	2 (1,23%)	6 (3,70%)	0	8 (4,94%)
7	-/-	-/-	+/+	-/-	-/-	-/-	0	0	1 (0,62%)	1 (0,62%)
8	-/-	-/-	+/+	-/-	+/-	-/-	0	1 (0,62%)	1 (0,62%)	2 (1,23%)
9	-/-	+/-	-/-	-/-	-/-	-/-	9 (5,55%)	24 (14,81%)	7 (4,32%)	40 (24,69%)
10	-/-	+/-	+/-	-/-	-/-	-/-	3 (1,85%)	3 (1,85%)	2 (1,23%)	8 (4,94%)
11	-/-	+/-	+/-	-/-	+/-	-/-	2 (1,23%)	1 (0,62%)	1 (0,62%)	4 (2,47%)
12	-/-	+/+	-/-	-/-	-/-	-/-	3 (1,85%)	3 (1,85%)	0	6 (3,70%)
13	-/-	+/+	+/-	-/-	+/-	-/-	0	0	1 (0,62%)	1 (0,62%)
14	+/-	-/-	-/-	-/-	-/-	-/-	4 (2,47%)	2 (1,23%)	0	6 (3,70%)
15	+/-	-/-	+/-	-/-	-/-	-/-	0	0	1 (0,62%)	1 (0,62%)
16	+/-	-/-	-/-	+/-	-/-	-/-	1 (0,62%)	0	0	1 (0,62%)
							63 (38,89%)	77 (47,53%)	22 (13,58%)	162

References

- Álvarez S, Mesa MS, Bandrés F, Arroyo E. C282Y and H63D mutation frequencies in a population from central Spain. *Dis Markers*. 2001;17(2):111-4.
- Aisen, P. The G277S mutation in transferrin does not disturb function. *Br J Haematol*. 2003;121(4):674-5.
- Beutler E, Gelbart T, Lee P, Trevino R, Fernandez MA, Fairbanks VF. Molecular characterization of a case of atransferrinemia. *Blood*. 2000;96(13):4071-4.
- Camaschella C. Understanding iron homeostasis through genetic analysis of hemochromatosis and related disorders. *Blood*. 2005;106(12):3710-7.
- Delanghe J, Verstraelen H, Pynaert I, Debels L, Taes Y, Verhasselt B, De Henauw S, Temmerman M. Human transferrin G277S mutation and iron deficiency in pregnancy. *Br J Haematol*. 2006;132(2):249-50.
- Lee PL, Halloran C, Trevino R, Felitti V, Beutler E. Human transferrin G277S mutation: a risk factor for iron deficiency anaemia. *Br J Haematol*. 2001;115(2):329-33.
- Zaahl MG, Merryweather-Clarke AT, Kotze MJ, van der Merwe S, Warnich L, Robson KJ. Analysis of genes implicated in iron regulation in individuals presenting with primary iron overload. *Hum Genet*. 2004;115(5):409-17.

5. DISCUSIÓN

Discusión

El trabajo de investigación desarrollado en esta tesis doctoral ha permitido profundizar en el conocimiento de la asociación entre marcadores genéticos y los trastornos del metabolismo del hierro. En concreto, se ha estudiado la anemia ferropénica y se incluye también un trabajo sobre la hemocromatosis, dos trastornos metabólicos relacionados con la deficiencia y el exceso de hierro, respectivamente.

En este sentido, en primer lugar se diseñó una herramienta de genotipado (reacción de mini-secuenciación *SNaPshot*), para replicar en población española la asociación entre marcadores genéticos procedentes de la literatura (SNPs) y parámetros bioquímicos relacionados con el metabolismo del hierro (hemoglobina, ferritina, transferrina, volumen corpuscular medio, hierro sérico y el receptor 1 de la transferrina).

Desde una perspectiva poblacional, se comparó la distribución alélica de estos SNPs relacionados con la regulación del hierro y sus estados patológicos (hemocromatosis y anemia ferropénica), en población española y europea. Este enfoque poblacional no sólo proporciona información valiosa genético poblacional sino que abunda en la validez de estos marcadores con posible utilidad clínica.

Finalmente, en un contexto más clínico se estudió una muestra de mujeres en edad fértil, para profundizar en la asociación de los marcadores genéticos con los parámetros bioquímicos relacionados con los fenotipos clínicos de la anemia ferropénica (anémico, deficiente y normal). De esta forma, se identificaron SNPs que influyen significativamente en el desarrollo de la anemia ferropénica, siendo algunos factores de riesgo y otros factores de protección.

A continuación, en el apartado 5.1, se procede a discutir los resultados más relevantes de esta tesis doctoral en función a los objetivos planteados. Para finalizar la discusión, en el apartado 5.2 se tratarán cuestiones que no responden directamente a los objetivos de la presente tesis pero que están relacionadas transversalmente con los trabajos incluidos.

5.1. Discusión por objetivos

Objetivo 1. Diseñar y valorar una herramienta de genotipado (reacción de mini-secuenciación *SNaPshot*), para replicar la asociación entre marcadores genéticos y parámetros bioquímicos relacionados con el metabolismo del hierro

Para la obtención de datos experimentales de la presente tesis doctoral, especialmente en relación a la ferropenia, se desarrolló una herramienta de genotipado simple y robusta para analizar de forma simultánea 10 marcadores SNP, que presentaban alta heterocigosidad en poblaciones europeas. En el artículo 2.1 se muestra el proceso de diseño y optimización de la herramienta. Esta reacción multiplex de minisecuenciación se empleó posteriormente en los artículos 3.1, 4.1, 4.2, comprobándose la utilidad de la misma.

La reacción de minisecuenciación *SNaPshot* consta de una primera PCR donde se amplifican los fragmentos en los que están incluidos los SNPs que se quieren analizar (primera fase). Sobre estos fragmentos se lleva a cabo la minisecuenciación propiamente dicha (segunda fase), en la que unos oligonucleótidos funcionan a modo de sondas uniéndose a la parte inmediatamente anterior al SNP que se quiere analizar (Sobrinó y Carracedo, 2009). Por lo tanto la optimización de la técnica se basa en desarrollar de forma equilibrada una primera PCR *multiplex*, en diseñar una reacción de minisecuenciación eficiente y finalmente en un exhaustivo trabajo de acoplamiento entre las dos fases. Durante la optimización de la técnica se probaron cantidades variables de ADN (desde 400 a 2ng/ml). Los resultados con poca cantidad de ADN fueron consistentes y con picos de una altura suficiente para el umbral de calidad que se había fijado (picos con una altura mayor de 200 RFUs o unidades de fluorescencia relativa). La capacidad de obtener buenos resultados con concentraciones variables nos permite genotipar muestras sin que sea necesario la cuantificación del ADN, lo que ahorra un tiempo considerable en el proceso completo de análisis (Budowle et al. 2008). Se minimizó la presencia de alelos inespecíficos y se obtuvo un electroferograma equilibrado, probando distintas

concentraciones de oligonucleótidos y diferentes temperaturas de anillamiento y de ciclos de la primera PCR *multiplex*. El incremento en el número de ciclos en la PCR fue la modificación que mejores resultados generó de cara a eliminar los picos inespecíficos. Los amplicones de la PCR fueron equilibrados para obtener un buen resultado en la minisequenciación. También para la reacción de *SNaPshot* fue necesario experimentar con varias concentraciones de sondas con el fin de obtener unos picos equilibrados. La cola poli-GACT en 5' resultó ser más estable para las extensiones más grandes de las sondas. Como era de esperar, los tamaños detectados por el secuenciador automático variaron respecto del tamaño real (Filipini et al. 2007). Estas diferencias se deben principalmente al tamaño del fragmento y al fluorocromo empleado para marcar las sondas (o *primer* de extensión) (Budowle et al. 2008). Debido a estas fluctuaciones, que pueden comprometer un genotipado correcto (López-Parra et al. 2013), las sondas se diseñaron para tener al menos 5 nucleótidos de diferencia en el tamaño de los fragmentos analizados, a fin de evitar solapamientos. Para comprobar la correcta asignación alélica se secuenciaron los amplicones de la primera PCR. En todos los casos hubo concordancia en la asignación entre secuenciación y el *SNaPshot*. Además, el SNP rs1375515 se tipó de nuevo con la plataforma *Sequenom* del CEGEN (Centro Nacional de Genotipado), no hallándose ninguna inconsistencia. La intensidad de los diferentes fluorocromos también varió aún después de la corrección realizada mediante la matriz de colores, pero no fue un impedimento para equilibrar el perfil genético.

Por otro lado, todos los SNPs empleados se encontraban en equilibrio Hardy-Weinberg aún cuando se analizaban en los tres grupos fenotípicos por separado. Este test se ha recomendado como una medida de calidad en el contexto de los estudios de asociación, lo cual se debe a que aquellos SNPs que se alejen del equilibrio podrían estar mal tipados o podrían ser objeto de presión selectiva (Wittke-Thompson et al. 2005). Otra medida para asegurar la calidad de los resultados es suprimir del análisis aquellos marcadores que tengan un alelo con una frecuencia inferior a 0,01 (variable no polimórfica). En esta reacción de *SNaPshot* todos los marcadores resultaron ser polimórficos (SNPs). Los resultados mostraron, además, gran heterocigosidad media en los grupos fenotípicos y una

alta heterocigosidad para cada uno de los SNPs, salvo para los casos de rs1799852 y rs1800562, en los que no se observaron homocigotos mutantes.

Como forma preliminar de valorar la utilidad de los SNPs seleccionados, se planteó la comparación de frecuencias de los mismos entre los diferentes grupos fenotípicos de la anemia ferropénica (normal, deficiente y anémico). Así, se observaron diferencias significativas en las frecuencias genotípicas al comparar los grupos anémico y control, especialmente para los SNPs rs1375515, rs3811647, y rs2280673. Además el rs1800562, aún siendo el marcador con menos heterocigosidad, fue capaz también de diferenciar entre los grupos normal y anémico. La frecuencia más alta para el genotipo AA del SNP rs2280673 en el grupo sano se reduce de forma marcada en el grupo de deficiencia de hierro, alcanzando el valor más bajo en el grupo de las anémicas. Lo contrario ocurre con el genotipo AA del SNP rs3811647 (homocigoto mutante) que aumenta en el grupo anémico. En el caso de rs1375515, seleccionado a partir de una región que mostró una señal de ligamiento con los niveles del receptor 1 de la transferrina en el estudio de Remacha et al. (2006), se aprecia que el homocigoto mutante se distribuye preferentemente en el grupo de anémicas. Sin embargo las diferencias sólo fueron significativas al comparar el grupo anémico frente a los otros dos (posible efecto recesivo). Este SNP (rs1375515) parecía ser el más informativo al discriminar entre el grupo normal y el anémico. En relación al rs1800562 se aprecia cómo el alelo que representa la mutación C282Y, es más frecuente en normales que en el resto de los grupos. Este resultado es el esperado teniendo en cuenta que esta mutación en homocigosis presenta un riesgo muy elevado de generar hemocromatosis y por tanto se asocia a mayores niveles de hierro (Feder et al. 1996). Estos SNPs se discutirán en mayor profundidad en los apartados relacionados con los objetivos 4, 5 y 6. El resto de SNPs no mostraron diferencias significativas en sus distribuciones genotípicas en función de los grupos. Se pensó, en un principio, que esto se podía deber a una falta de asociación entre estos SNPs y el desarrollo de la anemia, o también a que los límites definidos para los grupos fenotípicos no fueran los adecuados (Salas y Carracedo 2007). Sin embargo, los resultados de los artículos 4.1 y 4.2 así como otros trabajos (Blanco-Rojo et al. 2011), confirmaron la relación entre la práctica mayoría de los SNPs del *SNaPshot*

y ciertos parámetros bioquímicos relacionados con el metabolismo del hierro (ver apartados de discusión por objetivos 4, 5 y 6).

Desde el punto de vista metodológico, podemos afirmar que esta herramienta es sencilla de aplicar, posee buena capacidad de detección y es económica en relación a otros métodos de genotipado. La reacción de minisecuenciación se emplea desde la década de los 90 (Sokolov 1990), pero el uso del genotipado por *SNaPshot* se extendió a partir de los años 2000. Desde entonces, es una herramienta que ha demostrado ser útil para analizar mutaciones con diferentes propósitos ya sean, clínicos (Bustamante-Aragonés et al. 2008), de identificación forense, para aplicaciones en farmacogenética (López-Parra et al. 2013), para predicción de caracteres fenotípicos visibles (Söchtig et al. 2015) y otros. En coherencia con nuestros resultados, se puede afirmar que, a pesar de convivir con otras técnicas masivas de análisis, el *SNaPshot*, no ha perdido vigencia pues su idoneidad radica en ciertas ventajas y en un empleo en circunstancias concretas: el análisis de un número reducido de SNPs ya validados o con valor diagnóstico. Entre las ventajas específicas de este *SNaPshot*, se pueden citar: un coste estimado por muestra de 20 euros frente a los 130 euros de *TaqMan*®, otra técnica muy empleada a la hora de tipar un número reducido de SNPs (López-Parra et al. 2013), un tiempo de análisis corto (menos de 48 horas desde la recepción de las muestras) y una tasa de error despreciable.

Finalmente, a tenor de los resultados generados por este *SNaPshot* en los artículos 2.1, 3.1, 4.1 y 4.2, se tiene una visión global a posteriori del éxito y las limitaciones de la selección de SNPs empleada. El *SNaPshot* ha demostrado ser útil en cuanto aporta información relevante en relación a la base genética de los trastornos del metabolismo del hierro. Así, en el conjunto de los artículos 2.1, 3.1, 4.1 y 4.2 en los que se emplea esta técnica, se demuestra como todos los SNPs, salvo el rs4820268, se asocian significativamente con al menos uno de los parámetros hematológicos y bioquímicos estudiados. De tal forma que hemos sido capaces de replicar la asociación de marcadores obtenidos de otros estudios (Remacha et al. 2006, Benyamin et al. 2009a y 2009b, Chambers et al. 2009, Ganesh et al. 2009, Soranzo et al. 2009, Lo et al. 2011, Benyamin et al. 2014), y

además nos ha permitido descubrir un nuevo SNP (rs1375515) con especial relevancia en el metabolismo del hierro, como se verá en los siguientes apartados. En el artículo 4.2 se muestra como otros SNPs complementan la información obtenida a partir del *decaplex*. No se descarta, por tanto, la posibilidad de modificar la reacción para aumentar el número de SNPs y eliminar aquellos que no tengan especial relevancia o cuya información resulte redundante.

Objetivo 2. Comparar la variabilidad de los marcadores incluidos en el *SNaPshot* entre población española y otras caucásicas. Estudiar su posible estructura poblacional, con el fin de valorar la utilidad clínica de estas variantes

La consecución de este objetivo, se ve reflejada preferentemente en los artículos 2.1 y 3.1. La subestructuración de una población, en relación a ciertos marcadores, es uno de los motivos por los que se pueden encontrar asociaciones espurias entre estos marcadores y un fenotipo determinado (Balding 2006, Salas y Carracedo et al. 2007). Además, con el fin de estimar la validez en nuestra población, de los datos obtenidos del proyecto *HapMap*, se comparó la muestra estudiada en esta tesis doctoral con la de otras poblaciones europeas. Los datos que se obtuvieron del proyecto *HapMap* (valores de LD, información sobre bloques de haplotipos, etc.), fueron claves para desarrollar los análisis publicados en los artículos 4.1 y 4.2.

Las frecuencias de los genotipos de los tres grupos fenotípicos fueron comparadas locus por locus con los genotipos obtenidos de la población de origen europeo (colección CEPH), incluida en la base de datos del Proyecto International *HapMap* (International *HapMap* Consortium 2003). De forma general, los SNPs que se eligieron son variantes comunes en poblaciones europeas aunque hay ciertas excepciones. El SNP rs2673289, mostró diferencias significativas entre mujeres sanas y los datos del *HapMap*. Este resultado podía ser un efecto poblacional, siendo las frecuencias de este SNP típicas de población española. Para confirmar esta hipótesis habría que ampliar el estudio a otras poblaciones ibéricas. Además la

colección CEPH se compone de hombres y mujeres residentes en Utah (Estados Unidos), por tanto, no se puede descartar la segregación sexual de las frecuencias de este SNP ya que sólo se estudiaron mujeres. En cualquier caso esta diferencia no justificaría eliminar este SNP del *SNaPshot*, porque no se vería afectada su validez para población española. Además la comparación entre individuos no anémicos de la muestra e individuos no anémicos de la colección CEPH no pudo realizarse.

Las diferencias estadísticamente significativas de otros SNPs en los grupos deficientes y/o anémicos eran esperables ya que se eligieron precisamente por ser marcadores con un posible efecto en el estado del hierro y en parámetros relacionados. Este es el caso del rs1375515, que presenta diferencias significativas entre el grupo anémico y los datos de *HapMap*. Como se tratará en los apartados de la discusión por objetivos 4 y 6, la variante menos frecuente de este SNP y otros SNPs de la misma región cromosómica, se relacionan con mayor riesgo de pertenecer al grupo anémico. Esto explica que en el grupo anémico, estas variantes presentan frecuencias más altas que en la población europea del *HapMap* (CEPH) comparada. Este tag-SNP se seleccionó de una región que mostró una señal de ligamiento con los niveles del receptor 1 soluble de la transferrina en población española (Remacha et al. 2006). Este estudio se enmarcaba en el proyecto GAIT *Genetic Analysis of Idiopathic Thrombophilia* (Souto et al. 2003) y se discute la posible relación entre la enfermedad tromboembólica y la anemia ferropénica en el apartado del objetivo 4. También rs1800562 presentaba diferencias significativas entre el grupo deficiente y la población comparada de *HapMap* (CEPH). Con todo ello se confirma la eficiencia de estos SNPs, discriminando los diferentes grupos en relación a los fenotipos del estado del hierro en población europea. A la hora de testar la posible subestructuración de la muestra, no se observaron distancias genéticas (*FST*) significativas para los SNPs analizados entre los 3 grupos y/o entre las diferentes poblaciones caucasoides disponibles. Por lo tanto no se observa estratificación poblacional para estos marcadores ni en población caucasoide ni en la muestra española estudiada.

En resumen, a tenor de las diferentes distribuciones genotípicas encontradas entre los grupos y de la falta de estratificación poblacional, el

conjunto de SNPs, analizados mediante *SNaPshot*, se considera válido no solo para replicar los resultados significativos encontrados en los estudios de asociación genética relacionados con el metabolismo del hierro, sino también para encontrar nuevas asociaciones mediante SNPs no descritos con anterioridad.

Objetivo 3. Buscar variantes en los genes candidato del exceso de hierro y/o hemocromatosis (hepcidina y hemojuvelina) en individuos con diferentes genotipos para el gen *HFE*. Comparar su distribución alélica con la de otras poblaciones

Para dar respuesta a este objetivo en el artículo 3.2 se estudió una cohorte de 116 pacientes con sobrecarga de hierro agrupados de acuerdo a los genotipos de *HFE*. Se analizó en todas las muestras la región exónica del gen de la hepcidina (*HAMP*) y en dos grupos (sin mutaciones y con dobles heterocigotos C282Y/H63D) se secuenciaron los exones del gen de la hemojuvelina (*HJV*). Se analizaron también varios parámetros bioquímicos (hemoglobina, ferritina, volumen corpuscular medio, transferrina y la saturación de la transferrina). No se encontró ninguna mutación en los exones de los genes estudiados (*HJV*, *HAMP*) y tampoco se hallaron diferencias significativas en los valores de los parámetros en función de los distintos grupos según el genotipo de *HFE*. Estos resultados son llamativos si tenemos en cuenta que las mutaciones en *HFE* son la causa más frecuente (aproximadamente 80-90%) de la sobrecarga de hierro en población caucasoide (Feder et al. 1996, Niederau 2010). Este hecho apoya la posibilidad de que sean otros factores genéticos, no incluidos en *HJV* y *HAMP*, los causantes de la sobrecarga, especialmente en el grupo sin C282Y y en el grupo sin mutaciones. Cabe destacar a este respecto, que 30 individuos (25,9% de la cohorte) no presentaban ninguna mutación para cualquiera de los genes estudiados. Estos hallazgos sugieren que la hepcidina y la hemojuvelina están muy conservadas, al menos en las regiones exónicas, y que sus variantes son raras en poblaciones españolas, como ya ha sido señalado por otros autores, en individuos homocigotos para C282Y (Altès et al. 2009) y por nuestro propio grupo de investigación (Sarriá et al. 2007, Navas-Carretero et al. 2008, González et al. 2008). Además los

resultados parecen indicar que las regiones exónicas de *HJV* y *HAMP* no explican ni la diferente penetrancia de la hemocromatosis ligada a *HFE*, ni la etiología de otros tipos de sobrecarga de hierro.

La regulación de la mayoría de las rutas metabólicas del hierro se encuentra mediada en algún punto por la hepcidina (Abbaspour et al. 2014). Por su parte, la hemojuvelina ha sido descrita como un regulador clave de los niveles de hepcidina (Zhang 2010). De esto se infiere que el papel fundamental que juegan ambas proteínas haría que la presión selectiva actuara sobre ellas de modo que resultaran muy conservadas en términos evolutivos. Una revisión de la literatura confirma este razonamiento, ya que la mayoría mutaciones halladas se encuentran restringidas a determinadas familias (artículo 3.2). En esta línea, podemos observar que existe muy poca información, a nivel poblacional, de frecuencias alélicas para estas mutaciones. La mutación A197G en la hemojuvelina es relativamente frecuente, sin embargo, su implicación en la etiología de la sobrecarga de hierro se encuentra bajo discusión. De cualquier manera, no se debe descartar que otras mutaciones en regiones no codificantes de estos genes sean la verdadera causa de este trastorno (Bayele et al. 2006, Bayele y Srai 2009, Island et al. 2009), y que marcadores localizados en otros genes como la transferrina modifiquen su gravedad (de Tayrac et al. 2015).

Se cree que el origen de la mutación C282Y se encuentra en poblaciones celtas (Merryweather-Clarke et al. 2003), lo que se infiere por su mayor frecuencia en poblaciones del norte de Europa y por la clina descendente del noroeste al sureste de Europa (Campbell et al. 2003, Merryweather-Clarke et al. 2003). Concretamente en España, este patrón se mantiene al descender las frecuencias de norte a sur (Álvarez et al. 2001, Ropero-Gradilla et al. 2005). Así pues, tal y como demuestran nuestros datos, es esperable una mayor proporción de sobrecarga de hierro no ligada a *HFE* en poblaciones del sur de Europa. Asimismo, Radio et al. (2014) encontraron una proporción significativamente alta de hemocromatosis no ligada a *HFE* en una cohorte italiana. En ese caso se hallaron varias mutaciones en *TfR2* (receptor 2 de la transferrina).

En resumen, se puede afirmar que las regiones exónicas de *HJV* y *HAMP* raramente se encuentran involucradas en sobrecarga de hierro no ligada a *HFE*, y refuerzan la idea de que estas mutaciones son en su mayoría raras o exclusivas de una familia. Además, se confirma la menor prevalencia de sobrecarga de hierro y hemocromatosis ligada a *HFE* en poblaciones mediterráneas. Esto hace necesario estudiar en profundidad, empleando muestras de mayor tamaño, el origen de este desorden metabólico, especialmente en poblaciones con menor frecuencia de estas mutaciones, como la española.

Objetivo 4. Establecer asociaciones entre marcadores genéticos y parámetros bioquímicos y hematológicos relacionados con el metabolismo del hierro

Uno de los objetivos generales de la presente tesis doctoral es identificar marcadores genéticos relacionados con la anemia ferropénica. Este objetivo se aborda de una forma indirecta, al tratar de buscar asociaciones entre los marcadores y parámetros bioquímicos del metabolismo del hierro y, de forma directa, al relacionar estos marcadores con los fenotipos clínicos de la anemia. Estos objetivos se discuten en este apartado y en la discusión del objetivo 6, y se corresponden con los resultados obtenidos en los artículos 4.1 y 4.2. En este apartado se discutirán las asociaciones encontradas al analizar 45 SNPs con los parámetros bioquímicos y hematológicos. Estos resultados han proporcionado más información acerca de la regulación genética del metabolismo del hierro. Se procede a continuación a discutir sobre los SNPs que han resultado significativos. En el material suplementario del artículo 4.2 se contiene toda la información acerca de los 45 SNPs que fueron analizados. De éstos, 22 fueron asociados significativamente a algún parámetro bioquímico o hematológico.

Es conocido que los niveles séricos de transferrina correlacionan de forma inversa con otros parámetros bioquímicos ya que la expresión de la transferrina se regula de acuerdo a los requerimientos del organismo (Rolfs et al. 1997, Macedo et al. 2005). Esta es la razón por la cual los SNPs presentan efectos de signo contrario

para la transferrina en comparación con la hemoglobina, hematocrito, ferritina, hierro sérico y VCM (volumen corpuscular medio). El SNP rs1799852 (*TF*) es una excepción notable a esta norma ya que sus variantes menos frecuentes no sólo reducen los niveles de transferrina sino también los de hemoglobina y los de hematocrito. Este patrón singular de asociación induce a pensar que esta variante influye negativamente en el estado general del hierro, al reducir los niveles de expresión de la transferrina. Este SNP junto con rs3811647 y rs2280673 se localizan en el gen de la transferrina y se han asociado a los valores de esta proteína en sangre (Benyamin et al. 2009a, Kullo et al. 2010, Blanco-Rojo et al. 2011, Gichohi-Wainaina et al. 2015). Los resultados presentados en el artículo 4.2 replican estas asociaciones. Además, rs3811647 (*TF*) se ha relacionado con valores elevados de transferrina y le corresponde el nivel de asociación más significativo de los trabajos presentados en esta tesis. De forma general, niveles altos de transferrina son indicadores de un empeoramiento en el estado del hierro. Sin embargo, el rs3811647 se asoció exclusivamente a la transferrina lo que sugiere que este SNP influye en los niveles de esta proteína independientemente del estado del hierro. Por ello no se puede considerar un factor de riesgo para la ferropenia, lo que concuerda con los resultados de otros autores (McLaren et al. 2011, An et al. 2012).

Tal y como se ha observado en la literatura, el cromosoma 6 está muy relacionado con los parámetros implicados en la regulación del hierro (Ganesh et al. 2009, McLaren et al. 2011 y 2012). Este hecho se corrobora en esta tesis doctoral al encontrarse en este cromosoma la mayoría de los SNPs asociados (ver artículo 4.2). El SNP rs4895441 es una de las variantes intergénicas de *HBS1L-MYB* 6q23 asociada con niveles elevados de hemoglobina fetal (Ganesh et al. 2009, Kullo et al. 2010, Stadhouders et al. 2014, Mtatiro et al. 2015). El gen *MYB* se ha relacionado con la proliferación, supervivencia y diferenciación de las células progenitoras hematopoyéticas (Ganesh et al. 2009). La variante rara de este SNP genera un incremento en los valores de ferritina, hierro sérico y VCM, en concordancia con otros trabajos (Kullo et al. 2010). Este hecho explicaría por qué este SNP presenta un efecto protector en relación al estado del hierro, como se comentará en el siguiente apartado de la discusión, y como se ha observado en

Sankaran y Orkin (2013). Por otra parte, el SNP rs12216125 (6p22.2) se asocia a los valores de transferrina y VCM. Ganesch et al. (2009) lo relacionaron con VCM y hemoglobina, mientras que Benyamin et al. (2009b) con la saturación de la transferrina. No se ha propuesto aún ningún mecanismo subyacente, localizándose este SNP en una región intergénica en 6p22.2.

Otro marcador importante situado en el cromosoma 6 es la mutación C282Y (rs18000562), ya comentada en los apartados de discusión por objetivos 2 y 3. Localizada en el gen *HFE*, se encuentra asociada a rasgos eritrocitarios, especialmente hemoglobina y hematocrito, además de hierro sérico, la transferrina, y la saturación de la transferrina (Ganesh et al. 2009, Benyamin et al. 2009a, Tanaka et al. 2010, Lo et al. 2011, Oexle et al. 2011, McLaren et al. 2011, Abbaspour et al. 2014). La sustitución C282Y se relacionó con niveles altos de hemoglobina y ferritina, y se discute su papel como factor protector en el apartado de la discusión del objetivo 6. El efecto contrario se encuentra en la variante rara de rs1319441 en 6p21.33 ya que es un factor de riesgo de anemia ferropénica como se comentará al discutir el objetivo 6 (Ganesh et al. 2009, Benyamin et al. 2009b, Li et al. 2015). Localizado en *HIST1H2BJ*, un gen que codifica para una histona 1-J tipo H2B, no se conocen aún sus implicaciones fisiológicas.

Los SNPs rs855791 y rs4820268, localizados en el gen de la matriptasa 2 (*TMPRSS6*), se analizaron debido a su asociación con niveles bajos de hemoglobina, VCM, hierro sérico y con la deficiencia del hierro (Du et al. 2008, Benyamin et al. 2009a y 2009b, Chambers et al. 2009, An et al. 2012, Pei et al. 2014). También han sido asociados a otras enfermedades como la anemia refractaria al hierro (Melis et al. 2008, Heeney y Finberg 2014). En la presente tesis doctoral no se encontraron asociaciones significativas con ninguno de los parámetros estudiados, salvo con la transferrina en el caso de rs855791 (artículo 4.2). Una posible explicación es que los tamaños muestrales empleados no sean lo suficientemente grandes para revelar asociaciones con SNPs que tengan efectos pequeños. Los efectos de los SNPs que sí han resultado significativos parecen sensiblemente mayores que los efectos encontrados en otros trabajos para el SNP rs855791. Por ejemplo, el efecto de este SNP sobre la hemoglobina en Chambers et al. (2009) para el homocigoto

mutante era una reducción de 0,21g/dl en los niveles de esta proteína. En cambio para el homocigoto mutante de rs1375515, en el presente trabajo, la reducción fue de 0,54g/dl. El hecho de que al aumentar el tamaño muestral, el rs855791 apareciera asociado a la transferrina en el artículo 4.2, parece confirmar esa falta de potencia estadística para los efectos pequeños. Aunque no se replicaron los efectos adversos de rs855791 y rs4820268 sobre el estado del hierro, otro SNP, rs11089823, también localizado en *TMPRSS6* mostró relación con niveles altos de hemoglobina y hematocrito, tanto en la regresión lineal como en los árboles de decisión, y con niveles bajos de transferrina en el árbol de decisión (artículo 4.2). La proteína matriptasa 2 parece funcionar como sensor de los niveles de hierro, de tal forma que al reducirse éstos, la matriptasa 2 inhibe la expresión de la hepcidina (Du et al. 2008, Silvestri et al. 2008), movilizand las reservas del mineral y aumentándose su absorción en el intestino. Así, el rs11089823, podría estar incrementando la actividad de esta proteína, a juzgar por los efectos sobre hemoglobina, hematocrito y transferrina.

La región del cromosoma 2, donde se localizan los SNPs del *SNAPshot* rs2673289 y rs16826756, mostró la señal de ligamiento más fuerte con los niveles del receptor 1 de la transferrina en un estudio familiar (la otra señal, ligeramente más débil, se dio en una región del cromosoma 3 de la que se extrajo el rs1375515 en el mismo estudio familiar de Remacha et al. (2006). En el artículo 4.1, ninguno de estos dos SNPs del cromosoma 2 resultaron asociados a los parámetros analizados y se discutió que quizás esas variantes responsables de la señal eran específicas de ciertas familias. Sin embargo, en los análisis realizados en el trabajo 4.2, al aumentar el tamaño muestral de 284 a 358 mujeres, el rs16826756 se asoció con valores más bajos de VCM. Además el SNP rs2673289 presentó un modelo recesivo de efecto negativo sobre los niveles de hemoglobina, hematocrito y ferritina según el análisis de regresión lineal, y un efecto también negativo en los niveles de hemoglobina y hematocrito de acuerdo a los árboles de decisión. Por tanto, se confirma la conexión de esta región del cromosoma 2 con la regulación del metabolismo del hierro, en concreto con la ferropenia. En el trabajo de Remacha et al. (2006), que se encuadraba en el proyecto GAIT, se relacionaba la enfermedad tromboembólica con los niveles de receptor 1 de transferrina. Según

la hipótesis de Sullivan (1989), que otros trabajos han corroborado (Oboshi et al. 2015), a mayor grado de ferropenia existen menos probabilidades de sufrir trombosis. De lo que se deduce que estas variantes, al estar asociadas a una mayor probabilidad de ferropenia (niveles bajos de hemoglobina, ferritina y hematocrito), también estarían asociadas a una menor probabilidad de sufrir trombosis. La posible relación causal entre estos marcadores y los niveles del receptor 1 de la transferrina, deberá profundizarse en otros estudios.

En el artículo 4.1 se muestra por primera vez que el SNP rs1375515 (*CACNA2D3*) se asocia con parámetros bioquímicos relacionados con el hierro. Atendiendo a los resultados del BMA (*Bayesian Model Analysis*) este SNP se asoció con los niveles de VCM. También se asociaba en los árboles de decisión con ferritina y hemoglobina, que son los parámetros que emplea la OMS para definir los fenotipos clínicos del hierro (WHO 2011a, WHO 2011b). Los diferentes métodos estadísticos empleados revelaron que el comportamiento de rs1375515 es consistente en relación al efecto que parece tener sobre el estado del hierro. Así, el alelo G se relacionaba con niveles inferiores de ferritina (artículos 4.1 y 4.2), de hemoglobina (artículo 4.1) y de VCM (artículo 4.1). Los niveles bajos de estos parámetros son indicadores de riesgo de padecer anemia ferropénica. Por tanto, la presencia del alelo G se asocia significativamente con niveles bajos de estado de hierro. Además, estos resultados parecen apuntar a que el rs1375515 muestra un efecto recesivo en relación a la hemoglobina, ferritina y los fenotipos clínicos, ya que es el grupo homocigoto para el alelo menos frecuente (alelo G en este caso), el que presenta valores distintos con respecto a los demás (heterocigotos y homocigotos para alelo A). Los resultados de los modelos de regresión lineal son consistentes con la hipótesis del efecto recesivo. Dado que el test BMA sólo asume efectos codominantes, ésta podría ser la razón por la cual no se encontró para este test un efecto del SNP en hemoglobina y ferritina.

El SNP rs1375515 se eligió para el *SNaPshot* por ser un *tag*-SNP (según criterios de *HapMap*). Este marcador está localizado en una región intrónica de un gen que codifica una subunidad de un canal de calcio alfa voltaje dependiente (*CACNA2D3*). Las asociaciones significativas encontradas, por tanto, pueden

implicar que este SNP es el que ejerce influencia sobre los niveles de los parámetros o que, alternativamente, es otro SNP en desequilibrio de ligamiento (LD), el responsable de la relación estadística. El gen de la subunidad del canal del calcio comprende aproximadamente 1 Mbs del cromosoma 3, localizándose rs1375515 aproximadamente a 750kbs del extremo 5' y 300kbs del extremo 3' del gen (artículo 4.1). Los valores de LD muestran patrones distintos a lo largo de las poblaciones y descienden con la distancia genética, de tal forma que los bloques de LD de más de 100kbs son poco frecuentes (Daly et al. 2001). De acuerdo al análisis realizado con *Haploview* v4.1, rs1375515 se encuentra emplazado en un bloque de 25kbs de la región del canal. Si el responsable de la señal estadística fuera otro SNP en LD con éste, estaría también localizado muy probablemente en el mismo bloque o en otro adyacente y en cualquier caso dentro de la región del canal del calcio. Por tanto, atendiendo a las asociaciones encontradas para rs1375515, los datos de su localización apuntan a una posible relación entre las variantes del canal del calcio y parámetros bioquímicos empleados para valorar el estado del hierro (VCM, hemoglobina y ferritina).

En relación a la asociación con VCM encontrado en el artículo 4.1, y dada la localización en la subunidad del canal de calcio, existen numerosas evidencias que conectan el volumen celular y los canales iónicos con patologías del metabolismo del hierro. Así, la regulación precisa del volumen celular mediada por intercambio iónico es esencial para mantener una función celular normal, y en consecuencia cualquier alteración podría estar implicada en la fisiopatología de desórdenes tales como la insuficiencia hepática, fibrosis quística, y anemia falciforme (McManus et al. 1995, Lew et al. 2002, Tiffert et al. 2007). Además, los hematíes son muy sensibles a los cambios en volumen lo cual puede producir anemias hemolíticas (Lew et al. 2002). Sebastiani et al. (2010) descubrieron modificadores genéticos de la severidad de la anemia falciforme en un gen *KCNK6* que codifica para un canal de potasio. Aunque se desconocen los mecanismos fisiológicos de esta asociación, éste último trabajo parece apoyar el hecho de que el intercambio catiónico está de alguna forma relacionado con el volumen eritrocitario y por tanto con diferentes tipos de anemia.

Asimismo, los propios canales de calcio han sido asociados con el volumen celular y el estado del hierro. Lew et al. (2002) mostraron que bloquear los canales de calcio afectaba al volumen celular al causar deshidratación, generando anemia en última instancia. En esta línea, el bloqueo de los canales de calcio se ha postulado como tratamiento para cardiomiopatías por sobrecarga de hierro (Sugishita et al. 2009). Otros trabajos demuestran que los canales pueden ser directamente responsables de la toma de hierro como el canal tipo L dependiente de voltaje que podría ser una ruta alternativa para el hierro en su entrada en los cardiomiocitos (Gaasch et al. 2007). Por otra parte, su localización intrónica es consistente con su posible efecto ya que se han descrito numerosas variantes localizadas en intrones que afectan a la funcionalidad o a los niveles de expresión de la proteína en la que se encuentran (Benyamin et al. 2009a, 2009b y 2014, McLaren et al. 2011).

Estas evidencias parecen sugerir que la variante G del rs1375515, u otras variantes en LD, podría influir de alguna manera en la funcionalidad del canal del calcio pudiéndose incrementar el riesgo de sufrir anemia ferropénica. Los resultados del análisis genético en mayor profundidad (*Fine Mapping*) en relación a este canal se discuten en el siguiente apartado.

Objetivo 5. Estudiar las asociaciones obtenidas más relevantes, con el fin de proponer nuevas hipótesis sobre su implicación en la etiología de la anemia ferropénica

La novedad que suponía el hallazgo del rs1375515, y su posible relación con la ferropenia a través de su vínculo con los niveles bajos de hemoglobina, ferritina y VCM, hacían necesario explorar con mayor profundidad la región donde se encontraba. Además, este análisis más detallado podría permitir establecer una hipótesis más acertada sobre el papel fisiológico de la subunidad del canal de calcio (CACNA2D3), en relación al metabolismo del hierro. Así se realizó un “mapeo fino” (*Fine Mapping*) de la región genética que codifica para el canal (CACNA2D3). Este análisis apoya la hipótesis de que los canales de calcio

dependiente de voltaje podrían ser una vía directa de entrada del hierro a las células. Esta hipótesis ya enunciada en el artículo 4.1 se refuerza en el artículo 4.2, al hallarse varios SNPs de la región asociados con ferritina, hierro sérico, y los fenotipos clínicos.

Como se comentó en la introducción, un *Fine Mapping* es un análisis genético, en el que se genotipan varios SNPs de una región que anteriormente haya sido resultado asociado a algún fenotipo de interés. La selección de los SNPs se realiza teniendo en cuenta el LD de la región, de manera que se pueda captar la máxima variación tipando el menor número de SNPs posibles. Así el software empleado (*Haploview* v4.1), eligió un bloque de ligamiento de 100kb en el que rs1375515 estaba contenido. De este bloque seleccionó 16 SNPs, de los cuales 5 resultaron asociados con los niveles de hierro sérico y de ferritina (artículo 4.2).

La subunidad CACNA2D3 posee una función reguladora, modulando la densidad y la dinámica de activación/desactivación de los canales de calcio tipo L (Gurnett et al. 1996). Ya en el apartado anterior se discutió su asociación con el VCM en el marco de las evidencias que relacionan el intercambio iónico con el volumen celular y las patologías del hierro (Lew et al. 2002). Sin embargo, al realizar el *Fine Mapping* del gen *CACNA2D3* los resultados sugieren una relación más directa con el ión hierro Fe^{2+} (artículo 4.2), como se discutirá a continuación. De acuerdo con lo anterior, varios estudios han señalado cómo los canales de calcio voltaje dependientes pueden ser una ruta de entrada directa del hierro a los cardiomiocitos y las neuronas (Oudit et al. 2003, Gaash et al. 2007, Chattipakorn et al. 2011, Kumfu et al. 2012).

A tenor del patrón de LD de los SNPs del *Fine Mapping*, se observa cómo alguno de ellos no están en LD entre sí con lo que actuarían de forma independiente. Así, los efectos opuestos, y aparentemente independientes, de algunos SNPs de la región son consistentes con la función reguladora de CACNA2D3. Ejemplo de esto son rs1375515 y rs1868505 (en fuerte LD) que disminuyen los niveles de hemoglobina y ferritina, al contrario que rs49743366 y rs1868498 (en LD entre sí, pero no con los dos anteriores) los cuáles aumentan los

valores de hierro sérico y ferritina. Esto refuerza el papel potencial del canal como vía de entrada del hierro. Por añadidura, la ferritina no sólo es una proteína de almacenaje del hierro, sino que, como se ha demostrado en neuronas, puede actuar también como transportador extracelular de hierro y podría estar involucrada en la toma de hierro en una vía independiente de la transferrina (Chattipakorn et al. 2011). La ferritina se localiza de forma ubicua y su expresión se regula incrementándose a través de un elemento de respuesta al hierro (IRE, *Iron Response Element*). Además esta proteína está considerada como el parámetro más sensible y específico para identificar la deficiencia de hierro (Camaschella 2015).

Por otro lado, en condiciones de niveles de hierro elevados, el ión Fe^{2+} compite con el Ca^{2+} en la entrada al citoplasma a través de los canales de calcio provocando daños en cardiomiocitos y neuronas. Esta es la razón por la que el bloqueo de estos canales con ciertos fármacos se emplee como tratamiento de la sobrecarga de hierro (Sugishita et al. 2009), y pueda producir anemia (Lew et al. 2002).

La mayoría de los estudios anteriores que postulan los canales de calcio como vía entrada del hierro se han llevado a cabo con humanos con talasemia o sobrecarga, en cultivos celulares y/o en modelos murinos. Por tanto, en esta tesis doctoral se sugiere, por primera vez, que los canales de calcio tipo L regulados por CACNA2D3 podrían ser una vía directa para la entrada de hierro en individuos humanos sanos.

Objetivo 6. Establecer un modelo para identificar factores genéticos, de riesgo o protección, que influyen en el desarrollo de la anemia ferropénica

Como se ha comentado a lo largo de la discusión, varios estudios han relacionado marcadores genéticos con los parámetros hematológicos, pero pocos lo hacen específicamente con los fenotipos clínicos de la anemia ferropénica (McLaren et al. 2011, An et al. 2012, Blanco-Rojo et al. 2014). En los artículos que constituyen esta tesis doctoral se aborda la cuestión de la identificación de factores

de riesgo y protección de la ferropenia, mediante un árbol de decisión en el artículo 4.1, otro árbol con un número ampliado de SNPs en el artículo 4.2 (45 SNPs) y un modelo de regresión logística de pasos hacia atrás en el 4.2. Todos estos modelos aportaron información coherente y complementaria sobre la cuestión y han permitido replicar la asociación encontrada en otros trabajos publicados, además de revelar nuevas asociaciones con la anemia ferropénica.

En el artículo 4.1 el árbol de decisión fue obtenido considerando los fenotipos clínicos como variable dependiente y los 10 SNPs del *SNaPshot* como variables independientes. En este caso el árbol incluyó los SNPs rs1375515 (*CACNA2D3*) y rs1800562 (*HFE*). En el subgrupo de individuos que eran homocigotos para el alelo G de rs1375515, la proporción de anémicos casi se triplicaba en comparación con la muestra total (27,6% frente a 10,6%). Por otro lado, en aquellos individuos no homocigotos para el alelo G de rs1375515 y heterocigotos para rs1800562, la proporción de individuos normales casi se doblaba (77,8% frente a 43,7%), y desaparecían los anémicos. Se identificaban por tanto, un factor de riesgo y otro de protección, para la anemia ferropénica. De esta forma, el efecto combinado de rs1375515 y de C282Y, explicaba diferencias significativas en la distribución de los individuos entre los tres grupos fenotípicos. El árbol que se construyó en el artículo 4.2 con un mayor número de SNPs (45) y una muestra aumentada, confirma el factor de protección rs1800562, y descubre otro de riesgo, el SNP rs13194491 (*HBS1L*).

Por otra parte, en el artículo 4.2, analizando 45 SNPs, se muestra un modelo genético (regresión logística por pasos hacia atrás) que estima las probabilidades de pertenecer a un determinado grupo clínico. En este modelo se observa como cuatro SNPs explican una proporción significativa de la varianza del estado del hierro, de los cuales dos son factores protectores (rs1800562 en *HFE* y rs4895441 en *HBS1L*) y otros dos, rs13194491 (*HIST1H2BJ*) y rs1868505 (*CACNA2D3*), factores de riesgo de padecer anemia ferropénica. Este modelo representa un compendio coherente de los resultados obtenidos para los árboles de decisión comentados.

El modelo clasifica de forma correcta una proporción relativamente alta de individuos (45%) si nos atenemos a los valores de heredabilidad para hemoglobina y ferritina: 0,34-0,59 para hemoglobina y 0,20-0,29 para ferritina (Whitfield y Martin 1985, Garner et al. 2000, Souto et al. 2003, Sala et al. 2008). Sin embargo, dado que el grupo deficiente es el mayoritario, el porcentaje de acierto en la predicción proviene preferentemente de estimar los individuos deficientes como tales, lo que aumenta las reservas sobre la capacidad predictiva del modelo. Por tanto, una correcta lectura del modelo sería aquella en la que se relaciona a los marcadores incluidos, con una mayor o menor probabilidad de pertenecer a un determinado grupo fenotípico. De esta forma la probabilidad de ser anémico frente a deficiente y normal se incrementa en 1,69 veces para los portadores de AG de rs1868505 (*CACNA2D3*) y 3,07 veces para los portadores de CT en rs13194491 (*HIST1H2BJ*). Contrariamente, la probabilidad de ser normal frente a deficiente y anémico aumenta 14,93 veces en los portadores de AA para rs4895441 (*HBS1L*), lo que supone un efecto protector notable. De la misma forma, los heterocigotos GA de rs1800562 presentaban una probabilidad 6,06 veces mayor de ser normales frente a los demás fenotipos. La división del estado del hierro en dos grupos en vez de en tres, como aconseja el criterio de la OMS (WHO 2001), así como una muestra más grande, podrían mejorar las estimaciones del modelo, además de aumentar el número de SNPs incluidos.

Por último, cabe mencionar aquellos SNPs que, si bien no se asociaron directamente con los fenotipos clínicos, sí mostraron una gran importancia en relación a parámetros bioquímicos y hematológicos con lo que podrían ser relevantes en relación al estado del hierro. Este es el caso de rs11089823 (Matriptasa 2), rs1799852 (*TF*), y rs2673289 (cromosoma 2, señal de ligamiento con receptor soluble 1 de transferrina). Estos marcadores deben también considerarse factores de protección o de riesgo, si nos atenemos a los diferentes parámetros con los que se asocia, y a la alta significación. Finalmente, aunque el rs3811647 (*TF*) mostró la asociación más estadísticamente significativa, no se debe considerar un factor de riesgo ya que se relaciona exclusivamente con los niveles de transferrina sérica, sin que esto tenga, aparentemente, ningún efecto

sobre el estado general del hierro, en concordancia con otros trabajos (McLaren et al. 2011, An et al. 2012).

Es necesario recalcar que la muestra empleada posee características específicas que hacen que no sea una representación exacta de la población general. Los individuos son mujeres en edad fértil, ya que es el grupo poblacional con más riesgo de sufrir anemia. Atendiendo a los criterios de la OMS, dividimos la muestra en tres grupos clínicos, de los cuales el anémico y el de la deficiencia en hierro podrían estar sobrerrepresentados en relación a la población general. Aún así, a tenor del éxito a la hora de replicar la asociación de ciertos SNPs, no parece que la proporción de los grupos y el sexo estén siendo factores de confusión en estos análisis.

5.2. Temas transversales y consideraciones generales

En este último apartado de la discusión se abordan asuntos que, por su relación indirecta con los objetivos fundamentales de los artículos contenidos en la presente tesis doctoral, no han podido ser discutidos hasta ahora. Se ha considerado que estos temas poseen la suficiente importancia como para ser tratados en una sección independiente.

Papel de la selección natural en los marcadores genéticos

En relación a la selección natural, es plausible especular que aquellas variantes relacionadas con patologías o desórdenes metabólicos, como es el caso de la presente tesis, hayan experimentado presión selectiva en sus frecuencias alélicas. El efecto de esta presión parece revelarse en el análisis de los exones de hepcidina y hemojuvelina en relación con la hemocromatosis (artículo 3.2) donde no se encontró ninguna mutación. En relación a la anemia ferropénica se cumple además un patrón que es coherente con el efecto de la presión selectiva: aquellas variantes con un efecto fuerte sobre el parámetro con el que se asocian suelen

presentar frecuencias bajas, mientras que otras con efectos más tenues suelen ser variantes más comunes. Este fenómeno guarda cierta relación con los hallazgos presentados por el proyecto *HapMap* (*International HapMap Consortium 2007*). Éstos muestran cómo los SNPs no sinónimos presentan, de media, variantes más raras que los SNPs sinónimos. Este patrón tiene una excepción parcial que es el rs3811647. Este SNP muestra frecuencias relativamente equilibradas para sus tres genotipos, una significación muy alta en su asociación con la transferrina y un gran efecto en los niveles de esta proteína. Pero esta excepción es parcial en el sentido de que este SNP no se encuentra directamente asociado con la anemia ferropénica, aunque pueda ser considerada un factor de riesgo (Blanco-Rojo et al. 2011) y en otros casos parece que modera la gravedad de la hemocromatosis (de Tayrac et al. 2015). Tanto en el artículo 4.1 como en el 4.2 vemos como rs3811647 se asocia fuertemente con la transferrina pero no con la hemoglobina, con la ferritina, ni con la anemia ferropénica. De ahí, la posible falta de efecto de la presión selectiva para este marcador.

El gen del *HFE*, como ya se ha comentado anteriormente, es clave en el metabolismo del hierro. La mutación C282Y se ha revelado en esta tesis como un elemento fundamental en relación al estado del hierro (artículos 4.1 y 4.2). La baja frecuencia de homocigotos mutantes se explica por la severidad de la hemocromatosis. Sin embargo, la frecuencia de heterocigotos es más alta de lo que cabría pensar dada la gravedad de esta enfermedad y juega un papel protector como se observa en los artículos 4.1 y 4.2. En la línea de la hipótesis de la presión selectiva, las distribuciones actuales podrían haber sido fruto de dos efectos selectivos opuestos: el negativo por la patología de la hemocromatosis y la ventaja selectiva de poseer una variante que aumente significativamente los niveles de hierro en el organismo. Esta ventaja podría haber sido especialmente importante en condiciones nutricionales adversas, en las que la disponibilidad de hierro en la dieta fuera baja. Con todo, es difícil valorar, tanto en este caso como en otros, si la distribución no obedece meramente a dinámicas poblacionales en las que inciden movimientos migratorios, deriva génica o efecto fundador entre otros. Así, tal y como se explica en el artículo 3.2, la mutación *HFE* se originó en el norte de Europa, y su distribución de frecuencias posee un gradiente descendente hacia el

sur de Europa. Además, aún siendo la doble mutación C282Y potencialmente grave, su penetrancia se encuentra alrededor del 24 al 43% en hombres (Rossi et al. 2008), lo que justificaría una baja presión selectiva. En definitiva, si bien el efecto protector del genotipo heterocigoto es significativo e importante a nivel poblacional (art 4.1 y 4.2), no contamos con la suficiente evidencia como para concluir que haya habido una presión selectiva positiva sobre esta variante.

Estudio Genes-candidato y replicación de resultados de los estudios GWAS.

Una de las aproximaciones al problema de la herencia de los rasgos complejos es el estudio de los genes que codifican para proteínas implicadas en las rutas metabólicas relacionadas con un determinado rasgo, como es el caso del gen *HFE*. Esta aproximación se ve reforzada en ocasiones por los estudios familiares en los que se da a conocer una mutación con un efecto especialmente patológico. El abordaje gen-candidato ha revelado mutaciones que en su gran mayoría parecen ser muy raras o exclusivas de algunas familias. Sin embargo, esto provocaba que las variantes comunes con mayor implicación a nivel poblacional se mantuvieran sin conocer. Tal y como se demuestran en muchos estudios de asociación GWAS, son las variantes más frecuentes, las que contribuyen en mayor medida a explicar la base genética de las enfermedades comunes (Lohmueller et al. 2003). La línea seguida en la presente tesis doctoral se inició con el estudio complementario 4.3 y con el artículo 3.2. En ellos, al estudiar genes candidatos, se acabó concluyendo que estos genes, debido a su implicación clave en el metabolismo, suelen estar muy conservados y por tanto, son muy poco variables. Este es el motivo por el que se buscaron nuevos abordajes del problema, como el replicar en población española variantes significativas halladas en estudios de asociación GWAS. De esta manera, los estudios de asociación permiten detectar variantes en regiones y genes que no se hubieran planteado de forma inicial. Es a posteriori, una vez identificados las variantes con una posible relación con un fenotipo determinado, cuando se debe plantear los estudios más exhaustivos, en los genes o regiones que hayan resultado asociados. Ésta es la secuencia de pasos que se siguió en relación a la región

CACNA2D3, lo que permitió plantear una hipótesis sobre la posible implicación fisiológica de los canales de calcio en el metabolismo del hierro.

Análisis estadísticos, significación y tamaño muestral

Los métodos estadísticos empleados se discuten ligeramente en la discusión del objetivo 4, pero merecen una mención un tanto más extensa por su implicación en los resultados y conclusiones de la presente tesis doctoral. La combinación de diferentes análisis estadísticos, a riesgo de suponer un factor de confusión a la hora de interpretar los datos, ha resultado ser una forma productiva de extracción de información. Los diferentes métodos empleados en los trabajos han presentado poca disparidad en las conclusiones fundamentales y se han complementado en gran medida.

La estructura de los árboles de decisión depende tanto de la fuerza del efecto como de la frecuencia de la variante, y por supuesto de la competición con otros SNPs en la división de la muestra. Así, es posible que varios SNPs que tienen un efecto significativo por separado sobre la variable dependiente, no acaben reflejados en el árbol. Por ello, siempre se consideró necesario complementarlo con análisis más descriptivos como por ejemplo las regresiones lineales. El modelo de la regresión logística permitía valorar en el mismo plano el efecto conjunto de varios SNPs significativos sobre los fenotipos del estado del hierro. Ese es el motivo por el cual en el artículo 4.2 el árbol sólo presenta dos SNPs de los cuatro en la regresión logística. Esto también explica cómo muchos de los SNPs que resultan asociados en la regresión lineal no aparecen en los árboles de decisión.

El valor de los árboles de decisión radica en que su estructura permite analizar el efecto diferencial que puedan tener algunos SNPs, sobre un subgrupo determinado de la muestra que previamente haya sido definido por otro marcador. Con lo que es capaz de valorar el efecto conjunto de varios marcadores y sus posibles interacciones. Como ejemplo, en el árbol de la transferrina analizado en el artículo 4.2, se puede observar cómo en el subgrupo definido por los portadores de AA del SNP rs3811647 (*TF*), es donde el posible efecto protector del rs11089823

(*TMPRSS6*) es más significativo. Por tanto los árboles no sólo detectan efectos significativos sobre las variables, sino que nos informan de en qué segmentos de la muestra ese efecto es más importante.

Otra ventaja de los árboles de decisión es que estiman el posible modelo de efecto (dominante, recesivo o codominante), sin necesidad de establecer una hipótesis previa ya que estima cuál de las segregaciones de la muestra en función de los genotipos de los SNPs es la más significativa. Esta propiedad de los árboles también ha generado contrapartidas al generar en ocasiones, agrupaciones de la muestra espurias o sin ningún sentido biológico. Estos artefactos estadísticos se han corregido, ajustando los criterios de número mínimo de individuos en el nodo parental y filial, al tamaño muestral con el que se ha trabajado. En este sentido, una muestra de 358 individuos provocaba a veces que no todos los genotipos de los SNPs tuvieran un número suficiente para las tres clases (homocigoto normal, heterocigoto, homocigoto mutante). Por lo tanto, en ocasiones, el modelo de efecto estimado se debe interpretar con cautela, ya que podría variar al aumentar el tamaño muestral. En términos generales, este aumento debería producir árboles de decisión más estables que, además, representarían los efectos de los marcadores de forma más ajustada a la realidad.

Además, el tamaño muestral incide directamente en la significación, no sólo de los árboles de decisión, sino también del resto de los test estadísticos empleados. El hecho de que varios SNPs no se hayan podido replicar puede guardar relación con este aspecto de la muestra. En la discusión del objetivo 4 ya se ha tratado esta cuestión en relación al SNP rs855791. El no disponer de una muestra de mayor tamaño quizás esté provocando una pérdida de potencia estadística para los SNPs con efectos pequeños. En la terminología empleada en los estudios de asociación se llama “efecto *Jackpot*” al hecho de no poder replicar SNPs en otros estudios, posiblemente por errores de tipo II, causados por falta de potencia estadística (Chanock et. al 2007, Carracedo y Salas 2008). En esta línea, se observa como la ampliación de la muestra de 284 que se empleaba en el artículo 4.1 y en Blanco-Rojo et al. (2011), a 358 mujeres ocasionó que 4 SNPs más del *decaplex* resultaron asociados (artículo 4.2). Es de esperar, por tanto, que el

aumento de la muestra facilite el hallazgo de más SNPs relacionados significativamente con el estado del hierro. Por otro lado, otra de las causas de la falta de replicación es una mala definición del fenotipo de interés (Salas y Carracedo 2007). La división de los fenotipos clínicos que realiza la OMS y otros autores empleando hemoglobina y ferritina (WHO 2001, Navas-Carretero et al. 2008, WHO 2008, WHO 2011a, WHO 2011b) es una clasificación basada en la existencia de un grupo poblacional con deficiencia de hierro y sin anemia, con capacidades físicas y cognitivas mermadas. En este sentido, la diferencia en la gravedad de estos síntomas proporciona base clínica para separar deficientes de anémicos. Sin embargo, esto contrasta con la mayoría de los estudios de asociación que suelen establecer sólo dos categorías, a saber, sanos y enfermos. La división de los fenotipos de la anemia en dos grupos posiblemente facilitaría encontrar más SNPs asociados a este trastorno, especialmente cuando se emplean tamaños muestrales no muy elevados.

En cuanto a los criterios de significación que se han seguido, se debe resaltar que han sido considerablemente más bajos que los que se emplean en los estudios de asociación GWAS. En el caso de los estudios de asociación, la dificultad de replicar los resultados de un estudio en otras poblaciones independientes y la gran cantidad de SNPs e individuos valorados multiplica los análisis estadísticos de forma exponencial con lo que se aumentan las posibilidades de generar errores de tipo I, o hallazgos aleatorios. La tendencia de publicar valores muy bajos de significación para considerar un SNP como asociado (Ganesh et al. 2009, McLaren et al. 2011, Oexle et al. 2011, entre otros), posiblemente haya provocado que muchos estudios no pudiesen replicar ciertos hallazgos por falta de potencia estadística, como comentó Chanock et al. (2007). En los trabajos aquí presentados se emplea un nivel de significación del 0,05 que se considera adecuado en función del diseño experimental y el bajo número de SNPs valorados. Eso no es óbice para que a los árboles no se les aplique la corrección de Bonferroni, ya que cada uno de ellos genera varios análisis. Por otro lado, se da la circunstancia de que los SNPs han sido previamente seleccionados de otros estudios en los que ya resultaron significativamente asociados. Por tanto, considerando la evidencia a priori parece que no es necesario aumentar la exigencia de significación en nuestra muestra. Por

otro lado, la significación relativa de los SNPs en los estudios de asociación publicados en la literatura se mantiene, en términos generales, en los artículos incluidos en la presente tesis doctoral. Así los SNPs más significativos en la literatura también lo son en los artículos presentados, lo que refuerza la idea de que las relaciones encontradas no son espurias. Este hecho es de especial relevancia en aquellos marcadores que se testan por primera vez, como es el caso de los contenidos en el *Fine Mapping* (artículo 4.2). En conclusión, los resultados obtenidos refuerzan el posible acierto en los criterios de significación.

Consideraciones finales

Por último, cabe destacar que la presente tesis doctoral se centra en el análisis de la base genética de la anemia ferropénica, un trastorno en el que factores ambientales como la pérdida de sangre por menstruación, patologías o infecciones subyacentes y la nutrición también juegan un papel fundamental (Abbaspour et al. 2014, Blanco-Rojo et. al 2014). De hecho, se considera que en países desarrollados, la primera causa de anemia ferropénica es una dieta inadecuada en hierro (Camaschella 2015). A la luz de los resultados de la presente tesis doctoral, se refuerza el conocimiento de la base genética de la anemia ferropénica, hallándose factores genéticos que, presumiblemente interaccionando con los ambientales, pueden aumentar o disminuir el riesgo de desarrollar este desorden.

Como reflexión final, los hallazgos de la presente tesis doctoral han permitido profundizar en la asociación entre marcadores genéticos y los trastornos del metabolismo del hierro al haber: a) mostrado como la hepcidina y la hemojuvelina son genes altamente conservados en población española con sobrecarga de hierro, b) replicado, por primera vez, la asociación de determinados marcadores con parámetros relacionados con el estado del hierro y con la anemia ferropénica, en población española, c) identificado un nuevo SNP (rs1375515) relacionado con el estado del hierro, y d) proporcionado evidencia genética de que

los canales de calcio podrían ser una vía de entrada de hierro en condiciones fisiológicas.

6. CONCLUSIONES

Conclusiones

Conclusiones del objetivo 1. Diseñar y valorar una herramienta de genotipado (reacción de mini-secuenciación *SNaPshot*), para replicar la asociación entre marcadores genéticos y parámetros bioquímicos relacionados con el metabolismo del hierro

La reacción *multiplex* de minisecuenciación desarrollada es una herramienta sencilla, sensible, robusta, con una tasa de error despreciable y de bajo coste, lo que hace aconsejable su uso rutinario como técnica de genotipado y/o de diagnóstico clínico. Queda demostrada, también, su utilidad para la replicación de resultados de estudios de asociación.

La herramienta de genotipado diseñada (*SNaPshot*) es capaz de obtener resultados fiables trabajando con un rango variable de concentraciones del ADN molde (de 400 a 2 ng/ μ L).

El aumento de los ciclos en la primera PCR fue la modificación más efectiva de cara a minimizar la presencia de alelos inespecíficos y obtener un perfil equilibrado, en comparación con los ensayos a diferentes concentraciones de oligonucleótidos, temperaturas de anillamiento y ciclos de PCR.

Se demuestra que los criterios empleados para la selección de los SNPs incluidos en la reacción de minisecuenciación o *SNaPshot*, son apropiados, a tenor del conjunto de resultados obtenidos en esta tesis doctoral. Así, todos los marcadores, salvo el SNP rs4820268, se asociaron con al menos uno los parámetros bioquímicos estudiados.

Conclusiones del objetivo 2. Comparar la variabilidad de los marcadores incluidos en el *SNaPshot* entre población española y otras caucásicas. Estudiar su posible estructura poblacional, con el fin de valorar la utilidad clínica de estas variantes

Todos los marcadores incluidos en el *SNaPshot* pueden ser considerados variantes comunes tanto en población española como en otras poblaciones caucásicas. Únicamente el SNP rs2673289 mostró diferencias significativas entre las poblaciones europeas recogidas en *HapMap* y las mujeres normales (no anémicas) de nuestra muestra. El hecho de encontrar estas diferencias no invalida su empleo en población española.

Las diferentes distribuciones genotípicas encontradas entre los grupos fenotípicos/clínicos, la falta de estructura poblacional y los resultados de distancias genéticas hacen que el conjunto de SNPs analizados mediante *SNaPshot* sea de interés para replicar los resultados encontrados en los estudios de asociación genética relacionados con el metabolismo del hierro.

La semejanza entre nuestra muestra y otras poblaciones caucásicas, en cuanto a distancia genética se refiere, permite emplear aquellos tag-SNPs seleccionados a partir de la información de *HapMap* relativa al LD y a los posibles bloques de haplotipos. De esta forma, se refuerza la utilidad del *decaplex* desarrollado.

Conclusiones del objetivo 3. Buscar variantes en los genes candidato del exceso de hierro y/o hemocromatosis (hepcidina y hemojuvelina) en individuos con diferentes genotipos para el gen *HFE*. Comparar su distribución alélica con la de otras poblaciones

Las regiones exónicas de *HJV* y *HAMP* están raramente implicadas en la sobrecarga de hierro en población española, a juzgar por la ausencia de mutaciones encontrada en estos genes. Además, estas regiones están muy conservadas, según indica el análisis en nuestra muestra y la revisión de la literatura.

Se confirma una baja prevalencia de sobrecarga de hierro y hemocromatosis debida a *HFE* en poblaciones españolas, lo que concuerda con el gradiente noroeste-sureste que presenta la frecuencia de C282Y en Europa.

Conclusiones del objetivo 4. Establecer asociaciones entre marcadores genéticos y parámetros bioquímicos y hematológicos relacionados con el metabolismo del hierro

La variante menos frecuente del SNP rs3811647 (*TF*) está asociada exclusivamente, con una alta significación estadística, a niveles altos de transferrina sérica. No existe evidencia para considerar a este marcador como factor de riesgo para la anemia ferropénica de acuerdo a los resultados de la presente tesis doctoral.

El SNP rs1799852 (*TF*) se asocia con niveles bajos de transferrina y otros parámetros bioquímicos. Su localización sugiere que este SNP reduce la expresión de la transferrina sérica lo que genera un empeoramiento en el estado general del hierro.

La mutación C282Y (*HFE*) se relaciona con niveles más altos de hemoglobina y ferritina, y con niveles más bajos de transferrina.

La variante menos frecuente del SNP rs4895441, relacionado con la hemoglobina fetal y localizada en la región intergénica *HBS1L-MYB*, se asocia con niveles más altos de ferritina, hierro sérico y volumen corpuscular medio, y con niveles más bajos de transferrina.

La variante menos frecuente del SNP rs13194491 (*HIST1H2BJ*), se asocia con niveles más bajos de hemoglobina, hematocrito y hierro sérico.

El marcador rs11089823 se asocia a niveles altos de hemoglobina y hematocrito, se confirma así la implicación del gen de la matriptasa 2 (*TMPRSS6*) en el metabolismo del hierro.

Los SNPs rs2673289 y rs16826756 (cromosoma 2) replican la señal que relacionaba el riesgo de sufrir trombosis con los niveles del receptor 1 de la transferrina. De esta forma, se refuerza la asociación entre la anemia y la enfermedad tromboembólica.

El SNP rs1375515 está asociado con valores bajos de ferritina, de hemoglobina y de volumen corpuscular medio. La localización de este SNP en una región cromosómica que codifica una subunidad de un canal de calcio (*CACNA2D3*) sugiere que sus variantes afectan a la funcionalidad del canal.

Conclusiones del objetivo 5. Estudiar las asociaciones obtenidas más relevantes, con el fin de proponer nuevas hipótesis sobre su implicación en la etiología de la anemia ferropénica

Se refuerza el papel regulador de la subunidad alfa del canal de calcio a tenor de los efectos de diferente signo encontrados para los SNPs seleccionados mediante *Fine Mapping*.

Se proponen los canales de calcio tipo L como vía directa de entrada de hierro en condiciones fisiológicas, de acuerdo con las asociaciones entre los SNPs seleccionados mediante *Fine Mapping* y la ferritina y el hierro sérico. Esta hipótesis deberá ser explorada en futuros estudios de expresión, moleculares, modelos animales, y otros.

Conclusiones del objetivo 6. Establecer un modelo para identificar factores genéticos de riesgo o protección que influyen en el desarrollo de la anemia ferropénica

Dos variantes, el alelo A de rs1800562 en *HFE* (C282Y) y el G de rs4895441 en *HBS1L-MYB*, presentan un efecto protector en relación a la anemia ferropénica. Por el contrario, el alelo A de rs1868505 en *CACNA2D3* y el T de rs13194491 en *HIST1H2BJ* aumentan el riesgo de ser anémica en mujeres en edad reproductiva.

Otros SNPs que no se asociaron directamente con los fenotipos clínicos podrían considerarse factores de riesgo o protección de la anemia, a tenor de la relevancia de sus asociaciones con los parámetros del hierro. Así podrían ser factores de riesgo los SNPs rs2673289 (cromosoma 2) y rs1799852 (*TF*), y factor de protección el rs11089823 (*TMPRSS6*).

Conclusiones de los temas transversales y consideraciones generales

Se confirma el papel regulador del ADN intrónico e intergénico en el metabolismo del hierro ya que la mayoría de los SNPs para los que se ha hallado asociación se encuentran en regiones no exónicas.

El tamaño muestral empleado en estos estudios de asociación de marcadores genéticos con la anemia ferropénica, ha resultado ser, en general, suficiente para replicar los datos obtenidos de otros estudios de asociación. Sin embargo, la falta de asociación para el SNP rs855791 (*TMPRSS6*) parece indicar

una relativa falta de potencia estadística para los marcadores con efectos pequeños sobre el parámetro de interés.

La coherencia de nuestros resultados en comparación con otros trabajos, y la significación obtenida indican que no se han encontrado asociaciones espurias.

Los análisis de estudios multivariantes, tales como los árboles de decisión y la regresión logística por pasos hacia atrás, han aportado información coherente y complementaria sobre la relación de los marcadores genéticos con los trastornos del metabolismo del hierro.

Los árboles de decisión se han mostrado útiles a la hora de revelar posibles interacciones entre los factores genéticos que regulan el estado del hierro.

7. BIBLIOGRAFÍA

Bibliografía

A continuación se incluyen las referencias de la literatura citada en introducción (capítulo 1) y discusión (capítulo 6).

Abbaspour N, Hurrell R, Kelishadi R. Review on iron and its importance for human health. *J Res Med Sci.* 2014;19(2):164-74.

Altès A, Bach V, Ruiz A, Esteve A, Felez J, Remacha AF, Sardà MP, Baiget M. Mutations in HAMP and HJV genes and their impact on expression of clinical hemochromatosis in a cohort of 100 Spanish patients homozygous for the C282Y mutation of HFE gene. *Ann Hematol.* 2009;88(10):951-5.

Álvarez S, Mesa MS, Bandrés F, Arroyo E. C282Y and H63D mutation frequencies in a population from central Spain. *Dis Markers.* 2001;17(2):111-4.

An P, Wu Q, Wang H, Guan Y, Mu M, Liao Y, Zhou D, Song P, Wang C, Meng L, Man Q, Li L, Zhang J, Wang F. TMPRSS6, but not TF, TFR2 or BMP2 variants are associated with increased risk of iron-deficiency anemia. *Hum Mol Genet.* 2012;21(9):2124-31.

Anderson GJ, Darshan D, Wilkins SJ, Frazer DM. Regulation of systemic iron homeostasis: how the body responds to changes in iron demand. *Biometals.* 2007;20(3-4):665-74.

Bacon BR, Adams PC, Kowdley KV, Powell LW, Tavill AS; American Association for the Study of Liver Diseases. Diagnosis and management of hemochromatosis: 2011 practice guideline by the American Association for the Study of Liver Diseases. *Hepatology.* 2011;54(1):328-43.

Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet.* 2006;7(10):781-97.

Bayele HK, McArdle H, Srai SK. Cis and trans regulation of hepcidin expression by upstream stimulatory factor. *Blood.* 2006; 108:4237-45.

Bayele HK, Srai SK. Genetic Variation In Hepcidin Expression And Its Implications For Phenotypic Differences In Iron Metabolism. *Haematologica.* 2009;94:1185-8.

Beguín Y. Soluble transferrin receptor for the evaluation of erythropoiesis and iron status. *Clin Chim Acta.* 2003;329(1-2):9-22.

Benyamin B, Esko T, Ried JS, Radhakrishnan A, Vermeulen SH, Traglia M, Gögele M, Anderson D, Broer L, Podmore C, Luan J, Kutalik Z, Sanna S, van der Meer P, Tanaka T, et al. Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis. *Nat Commun.* 2014;5:4926.

- Benyamin B, McRae AF, Zhu G, Gordon S, Henders AK, Palotie A, Peltonen L, Martin NG, Montgomery GW, Whitfield JB, Visscher PM. Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels, *Am J Hum Genet.* 2009a;84(1):60-5.
- Benyamin B, Ferreira MA, Willemsen G, Gordon S, Middelberg RP, McEvoy BP, Hottenga JJ, Henders AK, Campbell MJ, Wallace L, Frazer IH, Heath AC, de Geus EJ, Nyholt DR, Visscher PM, Penninx BW, Boomsma DI, Martin NG, Montgomery GW, Whitfield JB. Common variants in TMPRSS6 are associated with iron status and erythrocyte volume. *Nat Genet* 2009b;41(11):1173-5.
- Beutler E, Van Geet C, te Loo DM, Gelbart T, Crain K, Truksa J, Lee PL. Polymorphisms and mutations of human TMPRSS6 in iron deficiency anemia. *Blood Cells Mol Dis.* 2010;44(1):16-21.
- Beutler E, Waalen J. The definition of anemia: what is the lower limit of normal of the blood hemoglobin concentration? *Blood.* 2006;107(5):1747-50.
- Blanco-Rojo R, Baeza-Richer C, López-Parra AM, Pérez-Granados AM, Brichs A, Bertoncini S, Buil A, Soria JM, Arroyo-Pardo E, Vaquero MP. Four variants in transferrin and HFE genes as potential markers of iron deficiency anaemia risk: an association study in menstruating women. *Nutr Metab (Lond).* 2011;8:69.
- Blanco-Rojo R, Toxqui L, López-Parra AM, Baeza-Richer C, Pérez-Granados AM, Arroyo-Pardo E, Vaquero MP. Influence of diet, menstruation and genetic factors on iron status: a cross-sectional study in Spanish women of childbearing age. *Neurobiol Aging.* 2014;35(4):777-85.
- Brakensiek K, Fegbeutel C, Mälzer M, Strüber M, Kreipe H, Stuhmann M. Juvenile hemochromatosis due to homozygosity for the G320V mutation in the HJV gene with fatal outcome. *Clin. Genet.* 2009;76:493-5.
- Budowle SA, Gonzalez S, Budowle B, Eisenberg AJ, Grange RW. A novel SNaPshot assay to detect the mdx mutation. *Muscle Nerve.* 2008;37(6):731-5.
- Bustamante-Aragones A, Gallego-Merlo J, Trujillo-Tiebas MJ, de Alba MR, Gonzalez-Gonzalez C, Glover G, Diego-Alvarez D, Ayuso C, Ramos C. New strategy for the prenatal detection / exclusion of paternal cystic fibrosis mutations in maternal plasma. *J Cyst Fibros.* 2008;7(6):505-10.
- Camaschella C, Roetto A, Calì A, De Gobbi M, Garozzo G, Carella M, Majorano N, Totaro A, Gasparini P. The gene TFR2 is mutated in a new type of haemochromatosis mapping to 7q22. *Nat Genet.* 2000;25(1):14-5.
- Camaschella C. Iron-Deficiency Anemia. *N Engl J Med.* 2015;373(5):485-6.
- Campbell S, George DK, Robb SD, Spooner R, McDonagh TA, Dargie HJ, Mills PR. The prevalence of haemochromatosis gene mutations in the West of Scotland and their relation to ischaemic heart disease. *Heart.* 2003;89:1023-6.

- Carracedo A, Salas A. 2008. Buscando el componente genético de las enfermedades complejas: estudios de asociación. En: *Genética Nutrición y Enfermedad*. pp. 103-12. EDIMSA (Eds). Madrid.
- Cazzola M, Cerani P, Rovati A, Iannone A, Claudiani G, Bergamaschi G. Juvenile genetic hemochromatosis is clinically and genetically distinct from the classical HLA-related disorder. *Blood*. 1998;92(8):2979-81.
- Chambers JC, Zhang WLI Y, Sehmi J, Wass MN, Zabaneh D, Hoggart C, Bayele H, McCarthy MI, Peltonen L, Freimer NB, Srai SK, Maxwell PH, Sternberg MJ, Ruokonen A, Abecasis G, Jarvelin MR, Scott J, Elliott P, Kooner JS. Genome-wide association study identifies variants in *TMPRSS6* associated with hemoglobin levels. *Nat Genet*. 2009;41(11):1170-2.
- Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P, Fraumeni JF Jr, Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover R, Kong CA, Merikangas KR, Morton CC, Palmer LJ, Phimister EG, Rice JP, Roberts J, Rotimi C, Tucker MA, Vogan KJ, Wacholder S, Wijsman EM, Winn DM, Collins FS. NCI-NHGRI Working Group on Replication in Association Studies. Replicating genotype-phenotype associations. *Nature*. 2007;447(7145):655-60.
- Chattipakorn N, Kumfu S, Fucharoen S, Chattipakorn S. Calcium channels and iron uptake into the heart. *World J Cardiol*. 2011;3(7):215-8.
- Cook JD. Diagnosis and management of iron-deficiency anaemia. *Best Pract Res Clin Haematol*. 2005;18(2):319-32.
- Core AB, Canali S, Babitt JL. Hemojuvelin and bone morphogenetic protein (BMP) signaling in iron homeostasis. *Front Pharmacol*. 2014;5:104.
- Crespo ÂC, Silva B, Marques L, Marcelino E, Maruta C, Costa S, Timóteo A, Vilares A, Couto FS, Faustino P, Correia AP, Verdelho A, Porto G, Guerreiro M, Herrero A, Costa C, de Mendonça A, Costa L, Martins M. Genetic and biochemical markers in patients with Alzheimer's disease support a concerted systemic iron homeostasis dysregulation. *Int J Mol Sci*. 2014;15(3):4077-87.
- Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES. High-resolution haplotype structure in the human genome. *Nat Genet*. 2001;29(2):229-32.
- Devalia V, Carter K, Walker AP, Perkins SJ, Worwood M, May A, Dooley JS. Autosomal dominant reticuloendothelial iron overload associated with a 3-base pair deletion in the ferroportin 1 gene (*SLC11A3*). *Blood*. 2002;100(2):695-7.
- Du X, She E, Gelbart T, Truksa J, Lee P, Xia Y, Khovananth K, Mudd S, Mann N, Moresco EM, Beutler E, Beutler B. The serine protease *TMPRSS6* is required to sense iron deficiency. *Science*. 2008;320(5879):1088-92.

- Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, Epstein CB, Frietze S, Harrow J, Kaul R, Khatun J, Lajoie BR, Landt SG, Lee BK, Pauli F, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57-74.
- Dunn LL, Suryo Rahmanto Y, Richardson DR. Iron uptake and metabolism in the new millennium. *Trends Cell Biol*. 2007;17(2):93-100.
- Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A, Dormishian F, Domingo R Jr, Ellis MC, Fullan A, Hinton LM, Jones NL, Kimmel BE, Kronmal GS, Lauer P, Lee VK, Loeb DB, Mapa FA, McClelland E, Meyer NC, Mintier GA, Moeller N, Moore T, Morikang E, Prass CE, Quintana L, Starnes SM, Schatzman RC, Brunke KJ, Drayna DT, Risch NJ, Bacon BR, Wolff RK. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet*. 1996;13(4):399-408.
- Filippini S, Blanco A, Fernández-Marmiesse A, Alvarez-Iglesias V, Ruíz-Ponte C, Carracedo A, Vega A. Multiplex SNaPshot for detection of BRCA1/2 common mutations in Spanish and Spanish related breast/ovarian cancer families. *BMC Med Genet*. 2007;8:40.
- Finberg KE. Iron-refractory iron deficiency anemia. *Semin Hematol*. 2009;46(4):378-86.
- Frenette PS, Atweh GF. Sickle cell disease: old discoveries, new concepts, and future promise. *J Clin Invest*. 2007;117(4): 850-8.
- Gaasch JA, Geldenhuys WJ, Lockman PR, Allen DD, Van der Schyf CJ. Voltage-gated Calcium Channels Provide an Alternate Route for Iron Uptake in Neuronal Cell Cultures. *Neurochem Res*. 2007;32(10):1686-93.
- Galesloot TE, Janss LL, Burgess S, Kiemeny LA, den Heijer M, de Graaf J, Holewijn S, Benyamin B, Whitfield JB, Swinkels DW, Vermeulen SH. Iron and hepcidin as risk factors in atherosclerosis: what do the genes say? *BMC Genet*. 2015;16:79.
- Ganesh SK, Zakai NA, van Rooij FJ, Soranzo N, Smith AV, Nalls MA, Chen MH, Kottgen A, Glazer NL, Dehghan A, Kuhnle B, Aspelund T, Yang Q, Tanaka T, Jaffe A, et al. Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nat Genet*. 2009;41(11):1191-8.
- Garner C, Tatu T, Reittie JE, Littlewood T, Darley J, Cervino S, Farrall M, Kelly P, Spector TD, Thein SL. Genetic influences on F cells and other hematologic variables: a twin heritability study. *Blood*. 2000;95(1):342-6.
- Gehrke SG, Pietrangelo A, Kascák M, Braner A, Eisold M, Kulaksiz H, Herrmann T, Hebling U, Bents K, Gugler R, Stremmel W. HJV gene mutations in European patients with juvenile hemochromatosis. *Clin Genet*. 2005;67(5):425-8.

- Gichohi-Wainaina WN, Melse-Boonstra A, Swinkels DW, Zimmermann MB, Feskens EJ1, Towers GW. Common Variants and Haplotypes in the TF, TNF-alpha, and TMPRSS6 Genes Are Associated with Iron Status in a Female Black South African Population. *J Nutr*. 2015;145(5):945-53.
- Gkouvatsos K, Papanikolaou G, Pantopoulos K Regulation of iron transport and the role of transferrin. *Biochim Biophys Acta*. 2012;1820(3):188-202.
- González FA, Ropero P, Peña A, Pérez C, Arroyo-Pardo E, Baeza Richer C, Villegas A. 2008. Papel de la hepcidina y la hemojuvelina en la hemocromatosis hereditaria. En: *Haematologica*, Spanish Edition, vol 93(2) pp. 60. Órgano Oficial de la Asociación Española de Hematología y Hemoterapia y de la Sociedad Española de Trombosis y Hemostasia. Acción Médica S.A. (Eds). Madrid, España.
- Gurnett CA, De Waard M, Campbell KP. Dual function of the voltage-dependent Ca²⁺ channel alpha 2 delta subunit in current stimulation and subunit interaction. *Neuron*. 1996;16(2):431-40.
- Hattori A, Wakusawa S, Hayashi H, Harashima A, Sanae F, Kawanaka M, Yamada G, Yano M, Yoshioka K. AVAQ 594-597 deletion of the TfR2 gene in a Japanese family with hemochromatosis. *Hepatol Res*. 2003;26(2):154-6.
- Heeney MM, Finberg KE. Iron-refractory iron deficiency anemia (IRIDA). *Hematol Oncol Clin North Am*. 2014;28(4):637-52.
- Hentze MW, Muckenthaler MU, Galy B, Camaschella C. Two to tango: regulation of Mammalian iron metabolism. *Cell*. 2010;142(1):24-38.
- Hentze MW, Muckenthaler MU, Andrews NC. Balancing acts: molecular control of mammalian iron metabolism. *Cell*. 2004;117(3):285-97.
- Huang FW, Rubio-Aliaga I, Kushner JP, Andrews NC & Fleming MD. Identification of a novel mutation (C321X) in HJV. *Blood*. 2004;104:2176-7.
- Hurrell R, Egli I. Iron bioavailability and dietary reference values. *Am J Clin Nutr*. 2010;91(5):1461S-7S.
- International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007;449(7164):851-61.
- International HapMap Consortium. The International HapMap Project. *Nature*. 2003;426(6968):789-96.
- Island ML, Jouanolle AM, Mosser A, Deugnier Y, David V, Brissot P, Loréal O. A new mutation in the hepcidin promoter impairs its BMP response and contributes to a severe phenotype in HFE related hemochromatosis. *Haematologica*. 2009;94:720-4.
- Jeffreys AJ, Kauppi L, Neumann R. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet*. 2001;29(2):217-22.

- Kullo IJ, Ding K, Jouni H, Smith CY, Chute CG. A genome-wide association study of red blood cell traits using the electronic medical record. *PLoS One*. 2010;5(9):e13011.
- Kumfu S, Chattipakorn S, Chinda K, Fucharoen S, Chattipakorn N. T-type calcium channel blockade improves survival and cardiovascular function in thalassemic mice. *Eur J Haematol*. 2012;88(6):535-48.
- Lanzara C, Roetto A, Daraio F, Rivard S, Ficarella R, Simard H, Cox TM, Cazzola M, Piperno A, Gimenez-Roqueplo AP, Grammatico P, Volinia S, Gasparini P, Camaschella C. Spectrum of hemojuvelin gene mutations in 1q-linked juvenile hemochromatosis. *Blood*. 2004;103:4317-21.
- Le Gac G, Férec C. The molecular genetics of haemochromatosis. *Eur J Hum Genet*. 2005;13(11):1172-85.
- Lee PL, Barton JC, Brandhagen D, Beutler E. Hemojuvelin (HJV) mutations in persons of European, African-American and Asian ancestry with adult onset haemochromatosis. *Br J Haematol*. 2004;127:224-9.
- Lew VL, Etzion Z, Bookchin RM. Dehydration response of sickle cells to sickling-induced Ca(++) permeabilization. *Blood*. 2002;99(7):2578-85.
- Lewontin RC, Kojima K. The Evolutionary Dynamics of Complex Polymorphisms. *Evolution*. 1960;14(4):458-72.
- Li J, Lange LA, Duan Q, Lu Y, Singleton AB, Zonderman AB, Evans MK, Li Y, Taylor HA, Willis MS, Nalls M, Wilson JG, Lange EM. Genome-wide admixture and association study of serum iron, ferritin, transferrin saturation and total iron binding capacity in African Americans. *Hum Mol Genet*. 2015;24(2):572-81.
- Lo KS, Wilson JG, Lange LA, Folsom AR, Galarneau G, Ganesh SK, Grant SF, Keating BJ, McCarroll SA, Mohler ER 3rd, O'Donnell CJ, Palmas W, Tang W, Tracy RP, Reiner AP, Lettre G. Genetic association analysis highlights new loci that modulate hematological trait variation in Caucasians and African Americans. *Hum Genet*. 2011;129(3):307-17.
- Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet*. 2003;33(2):177-82.
- López-Parra AM, Borobia AM, Baeza C, Arroyo-Pardo E, Carcas AJ. A multiplex assay to detect variations in the CYP2C9, VKORC1, CYP4F2 and APOE genes involved in acenocoumarol metabolism. *Clin Biochem*. 2013;46(1-2):167-9.
- Macedo MF, Cruz E, Lacerda R, Porto G, de Sousa M. Low serum transferrin levels in HFE C282Y homozygous subjects are associated with low CD8(+) T lymphocyte numbers. *Blood Cells Mol Dis*. 2005;35(3):319-25.

- Mattman A, Huntsman D, Lockitch G, Langlois S, Buskard N, Ralston D, Butterfield Y, Rodrigues P, Jones S, Porto G, Marra M, De Sousa M, Vatcher G. Transferrin receptor 2 (TfR2) and HFE mutational analysis in non-C282Y iron overload: identification of a novel TfR2 mutation. *Blood*. 2002;100(3):1075-7.
- Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutayavin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, Bates D, Hansen RS, Neph S, Sabo PJ, Heimfeld S, Raubitschek A, Ziegler S, Cotsapas C, Sotoodehnia N, Glass I, Sunyaev SR, Kaul R, Stamatoyannopoulos JA. Systematic localization of common disease-associated variation in regulatory DNA. *Science*. 2012;337(6099):1190-5.
- McLaren CE, Emond MJ, Subramaniam VN, Phatak PD, Barton JC, Adams PC, Goh JB, McDonald CJ, Powell LW, Gurrin LC, Allen KJ, Nickerson DA, Louie T, Ramm GA, Anderson GJ, McLaren GD. Exome sequencing in HFE C282Y homozygous men with extreme phenotypes identifies a GNPAT variant associated with severe iron overload. *Hepatology*. 2015;62(2):429-39.
- McLaren CE, Garner CP, Constantine CC, McLachlan S, Vulpe CD, Snively BM, Gordeuk VR, Nickerson DA, Cook JD, Leiendecker-Foster C, Beckman KB, Eckfeldt JH, Barcellos LF, Murray JA, Adams PC, Acton RT, Killeen AA, McLaren GD. Genome-wide association study identifies genetic loci associated with iron deficiency. *PLoS One*. 2011;6(3):e17390.
- McLaren CE, McLachlan S, Garner CP, Vulpe CD, Gordeuk VR, Eckfeldt JH, Adams PC, Acton RT, Murray JA, Leiendecker-Foster C, Snively BM, Barcellos LF, Cook JD, McLaren GD. Associations between single nucleotide polymorphisms in iron-related genes and iron status in multiethnic populations. *PLoS One*. 2012;7(6):e38339.
- McManus ML, Churchwell KB, Strange K. Regulation of cell volume in health and disease. *N Engl J Med*. 1995;333(19):1260-6.
- Melis MA, Cau M, Congiu R, Sole G, Barella S, Cao A, Westerman M, Cazzola M, Galanello R. A mutation in the TMPRSS6 gene, encoding a transmembrane serine protease that suppresses hepcidin production, in familial iron deficiency anemia refractory to oral iron. *Haematologica*. 2008;93(10):1473-9.
- Merryweather-Clarke AT, Cadet E, Bomford A, Capron D, Viprakasit V, Miller A, McHugh PJ, Chapman RW, Pointon JJ, Wilmhurst VL, Livesey KJ, Tanphaichitr V, Rochette J, Robson KJ. Digenic inheritance of mutations in HAMP and HFE results in different types of haemochromatosis. *Hum Mol Genet*. 2003;12:2241-7.
- Mtatiro SN, Mgaya J, Singh T, Mariki H, Rooks H, Soka D, Mmbando B, Thein SL, Barrett JC, Makani J, Cox SE, Menzel S. Genetic association of fetal-hemoglobin levels in individuals with sickle cell disease in Tanzania maps to conserved regulatory elements within the MYB core enhancer. *BMC Med Genet*. 2015;16:4.
- Muckenthaler MU. How mutant HFE causes hereditary hemochromatosis. *Blood*. 2014;124(8):1212-3.

- Murugan RC, Lee PL, Kalavar MR & Barton JC. Early age-of-onset iron overload and homozygosity for the novel hemojuvelin mutation HJV R54X (exon 3; c.160A-->T) in an African American male of West Indies descent. *Clin. Genet.* 2008;74:88-92.
- Nairz M, Weiss G. Molecular and clinical aspects of iron homeostasis: From anemia to hemochromatosis. *Wien Klin Wochenschr.* 2006;118(15-16):442-62.
- Navas-Carretero S, Baeza C, López-Parra AM, Tirado M, Arroyo-Pardo E, Vaquero MP. Ingesta de nutrientes y presencia de mutaciones en genes implicados en la deficiencia de hierro en un grupo de mujeres trabajadoras de la Comunidad de Madrid. *Nutr Hosp.* 2008;1(2):144-5.
- Niederau C. 2010. Hereditary Hemochromatosis and Iron Overload. En: *Clinical Hepatology: Principles and practices of hepatobiliary diseases.* pp. 1045-70. Dancygier H (Ed). Springer-Verlag. Berlin Heidelberg.
- Njajou OT, Alizadeh BZ, Aulchenko Y, Zillikens MC, Pols HA, Oostra BA, Swinkels DW, van Duijn CM. Heritability of serum iron, ferritin and transferrin saturation in a genetically isolated population, the Erasmus Rucphen Family (ERF) Study. *Hum Hered.* 2006;61(4):222-8.
- Nordborg M, Tavaré S. Linkage disequilibrium: what history has to tell us. *Trends Genet.* 2002;18(2):83-90.
- Oboshi M, Naito Y, Sawada H, Hirotani S, Iwasaku T, Okuhara Y, Morisawa D, Eguchi A, Nishimura K, Fujii K, Mano T, Ishihara M, Masuyama T. Temporary dietary iron restriction affects the process of thrombus resolution in a rat model of deep vein thrombosis. *PLoS One.* 2015;10(5):e0126611.
- Oexle K, Ried JS, Hicks AA, Tanaka T, Hayward C, Bruegel M, Gögele M, Lichtner P, Müller-Myhsok B, Döring A, Illig T, Schwienbacher C, Minelli C, Pichler I, Fiedler GM, Thiery J, Rudan I, Wright AF, Campbell H, Ferrucci L, Bandinelli S, Pramstaller PP, Wichmann HE, Gieger C, Winkelmann J, Meitinger T. Novel association to the proprotein convertase PCSK7 gene locus revealed by analysing soluble transferrin receptor (sTfR) levels. *Hum Mol Genet.* 2011;20(5):1042-7.
- OMIM, Online Mendelian Inheritance in Man. Atransferrinemia (Hypotransferrinemia, familial) #209300. [Último acceso, septiembre 2015]. Disponible en: <http://www.omim.org/>
- Orphanet, Portal de información de enfermedades raras y medicamentos huérfanos. Síndrome IRIDA (Anemia por deficiencia de hierro resistente al tratamiento por hierro) ORPHA209981. [Último acceso, septiembre de 2015]. Disponible en: <http://www.orpha.net/consor/cgi-bin/index.php>
- Ortiz I, Paredes Arquiola JM, López Serrano A, Moreno-Osset E. Hemocromatosis: etiopatogenia, diagnóstico y estrategia terapéutica. *Medicine.* 2012;11(19):1153-61.

- Oudit GY, Sun H, Trivieri MG, Koch SE, Dawood F, Ackerley C, Yazdanpanah M, Wilson GJ, Schwartz A, Liu PP, Backx PH. L-type Ca^{2+} channels provide a major pathway for iron entry into cardiomyocytes in iron-overload cardiomyopathy. *Nat Med*. 2003;9(9):1187-94.
- Papanikolaou G, Samuels ME, Ludwig EH, MacDonald ML, Franchini PL, Dubé MP, Andres L, MacFarlane J, Sakellaropoulos N, Politou M, Nemeth E, Thompson J, Risler JK, Zaborowska C, Babakaiff R, Radomski CC, Pape TD, Davidas O, Christakis J, Brissot P, Lockitch G, Ganz T, Hayden MR, Goldberg YP. Mutations in HFE2 cause iron overload in chromosome 1q-linked juvenile hemochromatosis. *Nat Genet*. 2004;36(1):77-82.
- Park CH, Valore EV, Waring AJ, Ganz T. Hepsidin, a urinary antimicrobial peptide synthesized in the liver. *J Biol Chem*. 2001;276(11):7806-10.
- Pei SN, Ma MC, You H, Fu HC, Kuo CY, Rau KM, Wang MC, Lee CT. TMPRSS6 rs855791 polymorphism influences the susceptibility to iron deficiency anemia in women at reproductive age. *Int J Med Sci*. 2014;11(6):614-9.
- Pichler I, Minelli C, Sanna S, Tanaka T, Schwienbacher C, Naitza S, Porcu E, Pattaro C, Busonero F, Zanon A, Maschio A, Melville SA, Grazia Piras M, Longo DL, Guralnik J, Hernandez D, Bandinelli S, Aigner E, Murphy AT, Wroblewski V, Marroni F, Theurl I, Gnewuch C, Schadt E, Mitterer M, Schlessinger D, Ferrucci L, Witcher DR, Hicks AA, Weiss G, Uda M, Pramstaller PP. Identification of a common variant in the TFR2 gene implicated in the physiological regulation of serum iron levels. *Hum Mol Genet*. 2011;20(6):1232-40.
- Pietrangelo A. Hereditary hemochromatosis--a new look at an old disease. *N Engl J Med*. 2004;350(23):2383-97.
- Pietrangelo A. Non-HFE hemochromatosis. *Semin Liver Dis*. 2005;25(4):450-60.
- Radio FC, Majore S, Binni F, Valiante M, Ricerca BM, De Bernardo C, Morrone A, Grammatico P. TFR2-related hereditary hemochromatosis as a frequent cause of primary iron overload in patients from Central-Southern Italy. *Blood Cells Mol Dis*. 2014;52:83-7.
- Ramírez L, Egaña B. 2003. Guía de conceptos de genética cuantitativa. Departamento de Producción Agraria, Universidad Pública de Navarra (Ed). Navarra, España. 12 pp.
- Ramsay AJ, Quesada V, Sanchez M, Garabaya C, Sardà MP, Baiget M, Remacha A, Velasco G, López-Otín C. Matriptase-2 mutations in iron-refractory iron deficiency anemia patients provide new insights into protease activation mechanisms. *Hum Mol Genet*. 2009;18(19):3673-83.
- Remacha AF, Souto JC, Soria JM, Buil A, Sarda MP, Lathrop M, Blangero J, Almasy L, Fontcuberta J. Genomewide linkage analysis of soluble transferrin receptor plasma levels. *Ann. Hematol*. 2006;85(1):25-8.

- Rivard SR, Lanzara C, Grimard D, Carella M, Simard H, Ficarella R, Simard R, D'Adamo AP, De Braekeleer M, Gasparini P. Autosomal dominant reticuloendothelial iron overload (HFE type 4) due to a new missense mutation in the FERROPORTIN 1 gene (SLC11A3) in a large French-Canadian family. *Haematologica*. 2003;88(7):824-6.
- Roetto A, Merryweather-Clarke AT, Daraio F, Livesey K, Pointon JJ, Barbabietola G, Piga A, Mackie PH, Robson KJ, Camaschella C. A valine deletion of ferroportin 1: a common mutation in hemochromatosis type 4. *Blood*. 2002;100(2):733-4.
- Roetto A, Papanikolaou G, Politou M, Alberti F, Girelli D, Christakis J, Loukopoulos D & Camaschella C. Mutant antimicrobial peptide hepcidin is associated with severe juvenile hemochromatosis. *Nat Genet*. 2003;33:21-2.
- Rolfs A, Kvietikova I, Gassmann M, Wenger RH. Oxygen-regulated transferrin expression is mediated by hypoxia-inducible factor-1. *J Biol Chem*. 1997;272(32):20055-62.
- Ropero-Gradilla P, González-Fernández FA, Briceño-Polacre O, Villegas-Martínez A. Geographical distribution of HFE C282Y and H63D mutation in Spain. *Med Clin (Barc)*. 2005;125:794-5.
- Rossi E, Olynyk JK, Jeffrey GP. Clinical penetrance of C282Y homozygous HFE hemochromatosis. *Expert Rev Hematol*. 2008;1(2):205-16.
- Ruiz Argüelles GJ, Ruiz Delgado GJ. 2014. *Fundamentos de Hematología* (5ª edición). Médica Panamericana (Ed). Madrid, España. 344 pp.
- Sala C, Ciullo M, Lanzara C, Nutile T, Bione S, Massacane R, d'Adamo P, Gasparini P, Toniolo D, Camaschella C. Variation of hemoglobin levels in normal Italian populations from genetic isolates, *Haematologica* 2008; 93(9):1372-5.
- Salas A, Carracedo A. Estudios de asociación en enfermedades complejas: problemas estadísticos relacionados con el análisis de polimorfismos genéticos. *Rev Clin Esp*. 2007;11:563-5.
- Sankaran VG, Orkin SH. The switch from fetal to adult hemoglobin. *Cold Spring Harb Perspect Med*. 2013;3(1):a011643.
- Sarriá B, López Parra AM, Navas-Carretero S, Pérez Granados AM, Baeza Richer C, Arroyo Pardo E, Vaquero MP. Hepcidin, transferrin (exon 7), and hemochromatosis genotyping suggests that haplotype block analysis is the best strategy for predicting iron deficiency phenotype in women. *Nutr Res*. 2007;27:672-8.
- Sebastiani P, Solovieff N, Hartley SW, Milton JN, Riva A, Dworkis DA, Melista E, Klings ES, Garrett ME, Telen MJ, Ashley-Koch A, Baldwin CT, Steinberg MH. Genetic modifiers of the severity of sickle cell anemia identified through a genome-wide association study. *Am J Hematol*. 2010;85(1):29-35.
- Silvestri L, Nai A, Pagani A, Camaschella C. The extrahepatic role of TFR2 in iron homeostasis. *Front Pharmacol*. 2014;5:93.

- Silvestri L, Pagani A, Nai A, De Domenico I, Kaplan J, Camaschella C. The serine protease matriptase-2 (TMPRSS6) inhibits hepcidin activation by cleaving membrane hemojuvelin. *Cell Metab.* 2008;8(6):502-11.
- Slatkin M. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat Rev Genet.* 2008;9(6):477-85.
- Smith JG, Newton-Cheh C. 2009. Candidate Gene and Genome-Wide Association Studies. En: *Cardiovascular Genetics and Genomics*. pp. 3-19. Dan Roden (Eds). Wiley-Blackwell. Oxford, UK.
- Sobrinho B, Carracedo A. 2009. SNP Typing in Forensic Genetics. En: *Forensic DNA Typing Protocols*. Carracedo A (Ed). Series: *Methods in Molecular Biology*, vol 297. pp 107-26. Humana Press Inc. Totowa, New Jersey, USA.
- Söchtig J, Phillips C, Maroñas O, Gómez-Tato A, Cruz R, Alvarez-Dios J, de Cal MÁ, Ruiz Y, Reich K, Fondevila M, Carracedo Á, Lareu MV. Exploration of SNP variants affecting hair colour prediction in Europeans. *Int J Legal Med.* 2015;129(5):963-75.
- Sokolov BP. Primer extension technique for the detection of single nucleotide in genomic DNA. *Nucleic Acids Res.* 1990;18(12):3671.
- Soranzo N, Spector TD, Mangino M, Kühnel B, Rendon A, Teumer A, Willenborg C, Wright B, Chen L, Li M, Salo P, Voight BF, Burns P, Laskowski RA, Xue Y, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat Genet.* 2009;41(11):1182-90.
- Souto JC, Remacha A, Buil A, Almasy L, Blangero J, Fontcuberta J. Genetic determinants of iron metabolism plasma phenotypes and their relationship with risk of thrombosis. *Haematologica.* 2003;88(12):1436-8.
- Stadhouders R, Aktuna S, Thongjuea S, Aghajani-refah A, Pourfarzad F, van Ijcken W, Lenhard B, Rooks H, Best S, Menzel S, Grosveld F, Thein SL, Soler E. HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. *J Clin Invest.* 2014;124(4):1699-710.
- Sugishita K, Asakawa M, Usui S, Takahashi T. A case of iron overload cardiomyopathy: beneficial effects of iron chelating agent and calcium channel blocker on left ventricular dysfunction. *Int Heart J.* 2009;50(6):829-38.
- Sullivan JL. The iron paradigm of ischemic heart disease. *Am Heart J.* 1989;117(5):1177-88.
- Tanaka T, Roy CN, Yao W, Matteini A, Semba RD, Arking D, Walston JD, Fried LP, Singleton A, Guralnik J, Abecasis GR, Bandinelli S, Longo DL, Ferrucci L. A genome-wide association analysis of serum iron concentrations. *Blood.* 2010;115(1):94-6.

- Tayrac de M, Roth MP, Jouanolle AM, Coppin H, le Gac G, Piperno A, Férec C, Pelucchi S, Scotet V, Bardou-Jacquet E, Ropert M, Bouvet R, Génin E, Mosser J, Deugnier Y. Genome-wide association study identifies TF as a significant modifier gene of iron metabolism in HFE hemochromatosis. *J Hepatol.* 2015;62(3):664-72.
- Tiffert T, Daw N, Etzion Z, Bookchin RM, Lew VL. Age decline in the activity of the Ca²⁺-sensitive K⁺ channel of human red blood cells. *J Gen Physiol.* 2007;129(5):429-36.
- Toxqui L, Vaquero MP. Chronic iron deficiency as an emerging risk factor for osteoporosis: a hypothesis. *Nutrients.* 2015;7(4):2324-44.
- Van der Harst P, Zhang W, Mateo Leach I, Rendon A, Verweij N, Sehmi J, Paul DS, Elling U, Allayee H, Li X, Radhakrishnan A, Tan ST, Voss K, Weichenberger CX, Albers CA, et al. Seventy-five genetic loci influencing the human red blood cell. *Nature.* 2012;492(7429):369-75.
- Wang CY, Meynard D, Lin HY. The role of TMPRSS6/matriptase-2 in iron regulation and anemia. *Front Pharmacol.* 2014;5:114.
- Weatherall DJ. Pathophysiology of thalassaemia. *Baillieres Clin Haematol.* 1998;11(1):127-46.
- Weiss G, Goodnough LT. Anemia of chronic disease. *N Engl J Med.* 2005;352(10):1011-23.
- Wells F. 2005. Haematinic Investigations. En: *The Science of Laboratory Diagnosis* (2^a edición). pp 312-22. Burnett D y Crocker J (Eds). Wiley & Sons, Ltd. West Sussex, UK.
- Whitfield JB, Martin NG. Genetic and environmental influences on the size and number of cells in the blood. *Genet Epidemiol.* 1985;2(2):133-44.
- WHO 2001. Iron Deficiency Anaemia Assessment, Prevention, and Control. A guide for programme managers. World Health Organization (Eds). Ginebra, Suiza. 140 pp.
- WHO 2008. Worldwide prevalence of anaemia 1993-2005. Global Database on Anaemia. de Benoist B, McLean E, Egli I, Cogswell M (Eds). World Health Organization. Ginebra, Suiza. 40 pp.
- WHO 2011a. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Vitamin and Mineral Nutrition Information System. World Health Organization (Eds). Ginebra, Suiza. 6 pp.
- WHO 2011b. Serum ferritin concentrations for the assessment of iron status and iron deficiency in populations. Vitamin and Mineral Nutrition Information System. World Health Organization (Eds). Ginebra, Suiza. 5 pp.
- WHO/CDC 2004. WHO/CDC expert consultation agrees on best indicators to assess iron deficiency, a major cause of anaemia. Report of a joint World Health Organization/ Centers for Disease Control and Prevention technical consultation on the assessment of iron status at the population level. Ginebra, Suiza. 2 pp.

- Wittke-Thompson JK, Pluzhnikov A, Cox NJ. Rational inferences about departures from Hardy-Weinberg equilibrium. *Am J Hum Genet.* 2005;76:967-86.
- Wood RJ, Ronnenberg AG. 2006. Iron. En: *Modern nutrition in health and disease* (10ª edición). pp 248-70. Shils ME, Shike M, Ross AC, Caballero B, Cousins RJ (Eds). Lippincott Williams & Wilkins. Philadelphia, USA.
- Zarychanski R, Houston DS. Anemia of chronic disease: a harmful disorder or an adaptive, beneficial response? *CMAJ.* 2008;179(4):333-7.
- Zhang AS, Anderson SA, Meyers KR, Hernandez C, Eisenstein RS, Enns CA. Evidence that inhibition of hemojuvelin shedding in response to iron is mediated through neogenin. *J Biol Chem.* 2007;282(17):12547-56.
- Zhang AS, Enns CA. Iron Homeostasis: Recently Identified Proteins Provide Insight into Novel Control Mechanisms. *J Biol Chem.* 2009;284(2): 711-5.
- Zhang AS. Control of systemic iron homeostasis by the hemojuvelin-hepcidin axis. *Adv Nutr.* 2010;1(1):38-45.

